Theory of the electronic states of semiconductor heterostructures

Coles, Richard A.

Theory of the Electronic States of Semiconductor Heterostructures

Richard A. Coles

A thesis submitted for the
degree of Doctor of Philosophy
at the University of Durham,
Department of Physics

March 1999

Abstract

This thesis is concerned with theoretical calculations of the properties of electronic bound states in semiconductor heterostructures. The complex band structure empirical pseudopotential method (EPM) is used as the foundation of the work. Spin orbit coupling and strain effects (due to lattice mismatch) are included in familiar ways, as is the transfer matrix method, allowing the study of arbitrarily configured heterostructures. These techniques are used to investigate the unusually deep InAs/AlSb conduction band well. The strong possibility of intraband transitions at electromagnetic wavelengths around 1.55μm is predicted, with corresponding enhanced momentum matrix elements and joint density of states over interband transitions. An InAs/GaSb/AlSb asymmetric well is investigated, paying particular attention to the bound states in the vicinity of the InAs/GaSb band overlap. The electron-like states are found to cross with heavy hole and anti-cross with light hole-like states, as a function of heterostructure dimension or applied electrostatic field. This is analogous to the hybridisation of states in the in-plane band structure, except that for zero in-plane wave vector there can be no appreciable hybridisation of electron and heavy hole states. A technique is described that has been developed to extract envelope functions from heterostructure wavefunctions calculated using the realistic complex band structure EPM approach. These envelope functions conform to Burt's theory (M. G. Burt, J. Phys.: Condens. Matt. **4**, 6651 (1992)) in that they are uniquely defined, continuous and smooth over all space. Comparisons with traditional effective mass envelope functions are made. The extracted envelope functions are used to demonstrate conclusively Burt's predictions (M. G. Burt, Superlatt. Microstruct. **17**, 335 (1995)) concerning the inadequacy of certain approximations for the calculation of interband dipole matrix elements and charge oscillation. Finally, the issue of $\mathbf{k} \cdot \mathbf{p}$ operator ordering is convincingly settled, in favour of 'ordered' over 'symmetrised' Hamiltonians, by comparison to EPM calculations, and using EPM derived $\mathbf{k} \cdot \mathbf{p}$ parameters.

Declaration

The work presented in this thesis has been carried out by the candidate (except where acknowledged otherwise) and has not been previously submitted for any degree.

_____

PhD Candidate

_____

PhD Supervisor

Acknowledgments

Firstly, special thanks go my PhD supervisors, Dick Abram and Stuart Brand, for their patient assistance and unwavering support throughout my period of study. I would also like to thank BT for providing extra funding in the form of a CASE award, and in particular Mike Burt for offering many helpful and interesting suggestions with such infectious enthusiasm. In addition, thanks go to EPSRC for providing the bulk of the funding for this research.

On the non-academic side, I would like to thank my family for their encouragement over the years, especially my Mum, Dad and Grandma, and my brother, Ed, who as an experienced undergraduate wasted little time in introducing me to some of the better public houses Durham has to offer. I would also like to thank Richard Coulam, without whose help and single-minded dedication during the dark days of undergraduate life in London I probably would not have had the opportunity to embark on a PhD course. Of those in Durham, my life has been made immeasurably more fun with the help of Josep Alabert, Suzanne Duxbury, Moncef Fakh Fakh, Yasu Kagawa, Andrew Pocklington, Uta Starcke, Hazel Tucker and especially Amanda Winn.

Finally, I would like to express my appreciation to all those in the Physics Department, academic and non-academic, who helped make my time here so enjoyable. In particular, I'd like to thank the members of the Condensed Matter Theory group, past and present, especially Chris "Topher" Caulfield, Stewart "Teetotal" Clark, Gavin "Pixie Chair" Crow, Dave "Passing Machine" Dugdale, Dan "Kobayashi Maru" Harrison, Dave "Constant Snacker" Hoare, Hervé "Err-Vay" Ness, Steve "Scary Shorts" Pugh and Des "Powerful Thews" Ryan. Their assistance was vital during important experimental work on the peculiar properties of fluids when tested in imperial measures! A big thanks to you all.

To Amanda

Contents

List of Figures

List of Tables

Symbols and acronyms

Below is a list of symbols commonly used in this text, together with a brief explanation of their meaning. Quantities in bold type-face are vectors.

Symbol	Meaning
$\uparrow$	Spin up
$\downarrow$	Spin down
$\times$	Scalar multiplication (*not* vector product)
$\wedge$	Vector (cross) product
$\parallel$	Indicates the *in-plane* direction, i.e., planes parallel to heterostructure interfaces (xy-plane in this work)
$\perp$	Indicates the *perpendicular* direction, i.e., the heterostructure growth direction (the z-direction in this work)
Δ_0	Spin orbit split-off energy
Γ	Reciprocal space Brillouin zone centre, $\mathbf{k} = (0, 0, 0)$
λ_a	Antisymmetric EPM spin parameter
λ_s	Symmetric EPM spin parameter
φ	Electron wavefunction
ψ	Electron wavefunction
Ψ	Electron wavefunction
σ_{ij}	Mechanical stress tensor
σ_x	x-component of Pauli spin matrix
σ_y	y-component of Pauli spin matrix
σ_z	z-component of Pauli spin matrix
$\boldsymbol{\sigma}$	Pauli spin matrix

Symbol	Meaning
$\boldsymbol{\tau}$	Position in the primitive unit cell of atom B in an AB zinc-blende material, with the origin midway between the two atoms, and atom A at $-\boldsymbol{\tau}$
$a_{\mathbf{g}}$	Fourier expansion coefficient for reciprocal lattice vector $\mathbf{g}$
$\mathbf{e}$	Unit vector
$\mathbf{e}_x$	Unit vector in x direction
$\mathbf{e}_y$	Unit vector in y direction
$\mathbf{e}_z$	Unit vector in z direction
E	Energy
$f_n(z)$	z-dependent part of the n^{th} band electron envelope function
$F_n(\mathbf{r})$	n^{th} band electron envelope function
$g^a_{\mathbf{g}}$	Gradient, in reciprocal space, of the antisymmetric pseudopotential Fourier coefficient at $\mathbf{g}$
$g^s_{\mathbf{g}}$	Gradient, in reciprocal space, of the symmetric pseudopotential Fourier coefficient at $\mathbf{g}$
g	Reciprocal lattice vector component along the well growth direction
$\mathbf{g}$	Reciprocal lattice vector
$\mathbf{g}_{\|\|}$	In-plane components of the reciprocal lattice vector, $\mathbf{g} = \mathbf{g}_{\|\|} + \mathbf{e}_z g$
h	Planck's constant
$\hbar$	$h/2\pi$
$\widehat{H}$	Hamiltonian operator
$\widehat{H}^p$	Pseudopotential Hamiltonian operator
k	Component of electron wave vector along heterostructure growth direction
$\mathbf{k}$	Electron wave vector
$\mathbf{k}_{\|\|}$	In-plane components of the electron wave vector, $\mathbf{k} = \mathbf{k}_{\|\|} + \mathbf{e}_z k$
k_b	Boltzmann's constant
K	Reciprocal space points of the type $\{\frac{3}{4}\frac{3}{4}0\}$
L	Reciprocal space points of the type $\{\frac{1}{2}\frac{1}{2}\frac{1}{2}\}$
m_e	Electron rest mass
$\mathbf{p}$	Momentum
$\widehat{\mathbf{p}}$	Momentum operator, $-i\hbar\nabla$
$\mathbf{q}$	Reciprocal space vector
$\mathbf{r}$	Real space position vector
$\mathbf{r}_{\|\|}$	In-plane components of the real space postition vector, $\mathbf{r} = \mathbf{r}_{\|\|} + \mathbf{e}_z z$
$s^a_{\mathbf{g}}$	Antisymmetric structure factor, $\sin(\mathbf{g} \cdot \boldsymbol{\tau})$

Symbol	Meaning
$s^s_{\mathbf{g}}$	Symmetric structure factor, $\cos(\mathbf{g} \cdot \boldsymbol{\tau})$
$u_n(\mathbf{r})$	Bloch periodic function for the n^{th} energy band
$u_{n\mathbf{g}}$	Fourier coefficient of $u_n(\mathbf{r})$, for reciprocal lattice vector $\mathbf{g}$
U	Reciprocal space points of the type $\{1\frac{1}{4}\frac{1}{4}\}$
$v^a_{\mathbf{g}}$	Antisymmetric EPM pseudopotential form factor (Fourier coefficient) for reciprocal lattice vector $\mathbf{g}$
$v^s_{\mathbf{g}}$	Symmetric EPM pseudopotential form factor (Fourier coefficient) for reciprocal lattice vector $\mathbf{g}$
V	Potential
V^p	Local pseudopotential
$\widehat{W}$	Pseudopotential operator
X	Reciprocal space points of the type $\{100\}$

Below is a table of acronyms used in this text, together with their meanings.

Acronym	Meaning
CBS	Complex Band Structure
DBE	Double Band Edge
EFA	Envelope Function Approximation
EM	Electromagnetic
EPM	Empirical Pseudopotential Method
LCAO	Linear Combination of Atomic Orbitals
SBE	Single Band Edge
SOC	Spin Orbit Coupling
SPB	Single Parabolic Band
SQW	Single Quantum Well
SSO	Spin orbit split-off

Chapter 1

Introduction

1.1 History and background

This thesis deals with the theory of semiconductor heterostructures, i.e., structures consisting of more than one semiconductor material, and specifically those with all the material interfaces parallel. When the interfaces, or *heterojunctions*, between the different materials, are separated by distances on a scale comparable to that of the electron's de Broglie wavelength, resulting in 'thin' layers, the system must be analysed quantum mechanically. Such structures allow the tailoring of their physical properties in much more varied ways than are possible using bulk materials and their alloys. They also lead to novel effects not seen in the bulk. In particular, this work is concerned with heterostructures which have energy ranges in which the electrons are spatially confined. A common example of a heterostructure is the single quantum well (SQW), the basic concept of which is familiar from simple textbook calculations (see, for example, p275 of [1]). Such textbook analysis of SQWs was usually carried out either as an approximation to conduction electrons in metals or molecules, those bound in the atom or as a purely quantum mechanical exercise. However, the rapid advancement over the last 20–30 years of material growth techniques has meant

that SQW and other heterostructure calculations now have direct applications to real systems.

Heterostructures of a type similar to those discussed in this work were first conceived by Esaki and Tsu in 1970 [2], who suggested the growth of alternating layers of two different semiconductor materials, with atomically abrupt interfaces between them, as a way of producing further 'materials' capable of exhibiting negative differential resistance[1] (NDR). Such a periodic structure is called a *superlattice*. At the time, though, growth techniques were not sufficiently developed to fully realise such a structure.

In the late 1960s, two methods with the potential to grow heterostructures of the desired scale were developed, namely molecular beam epitaxy (MBE) [4] and metal organic chemical vapour deposition (MOCVD) [5]. Over the years, these methods have been refined to the extent that, for many material systems, the fabrication of heterostructures with genuinely abrupt interfaces and thin layers can be considered as relatively routine (see, for example, [6,7]). Such heterostructures have found, or are promising for, applications in many areas, including lasers, transistors, tunelling diodes, photodetectors and non-linear optical devices (such as all optical switches and frequency convertors).

The advances in experimental technique have encouraged much effort in the theoretical investigation of heterostructures. Early calculations [8,9] used essentially textbook Kronig-Penney models (see, for example, p286 of [1]). The continually increasing computational power and relative affordability of digital computers has allowed more sophisiticated models to be adopted. These include various tight-binding methods [10–12], $\mathbf{k} \cdot \mathbf{p}$ and envelope function methods [13–18], and *ab initio* [19] and empirical [20,21] pseudopotential methods (see [22] for a detailed review of

[1] Of the bulk materials, only GaAs, InP and $In_xGa_{1-x}As$ satisfy the conditions for NDR to occur (see, for example, p218 of [3]). It should be noted that NDR in heterostructures is due to a different mechanism from that in bulk semiconductors.

various theoretical methods for studying superlattices). Of these methods, the six or eight band $\mathbf{k} \cdot \mathbf{p}$ method and envelope function approaches have been the most widely used, whilst at the same time, their range of applicability and mathematical foundation was until recently not well understood.

This thesis uses an empirical pseudopotential method to study the properties of bound electronic states in various semiconductor heterostructures, consisting of group III-V and II-VI compounds and alloys. In addition, the properties of envelope functions extracted from the microscopic pseudopotential calculations are examined and compared to envelope functions obtained in the traditional way. In particular, we focus on the issues surrounding the appropriate interface boundary conditions to be applied in effective mass envelope function approximations.

1.2 Thesis plan

Chapter 2

We first give a brief review of the concepts of pseudopotential theory, including the key equations, and then concentrate on how the general theory is reduced to the particular empirical pseudopotential method (EPM) used in this work. Next, the methods for including spin orbit coupling (SOC) and strain effects are introduced. The section on strain also discusses lattice mis-match and epitaxial growth, lattice mismatch being the main reason for including strain in the calculations. We then describe the procedures used to obtain the various parameters required by the EPM, followed by a discussion of the complex band structure technique, where one component of the electron wave vector is generalised to a complex number. This technique is used as the basis of the heterostructure calculations. The convergence of the calculated values with the number of terms in the Fourier expansion of the

wavefunction is briefly covered, and finally, examples of the calculation of various electronic properties of bulk crystals are presented.

Chapter 3

This chapter covers the theory and methods used in the EPM calculations of the electronic bound states of heterostructures considered in this thesis. After a brief introduction, the critical issue of the wavefunction interface boundary conditions is discussed, followed by a description of the transfer matrix method for wavefunction matching over many interfaces, as implemented in the complex band structure EPM scheme, and the equations used for determining the existence of bound states. The remainder of the chapter is devoted to the concepts and methods used for the calculation of various electronic properties of heterostructures, specifically, momentum and dipole matrix elements between bound states and the joint density of states. Finally, we consider the application of an electric field along the heterostructure growth direction.

Chapter 4

Chapter 4 presents the results of heterostructure calculations performed using the methods described in the preceding chapter. Two systems are investigated; an InAs/AlSb single quantum well (SQW) and an asymmetric InAs/GaSb/AlSb system. The InAs/AlSb system forms an unusually deep conduction band well, and the results focus on the possibilites of transitions between conduction band bound states (intraband transitions) at the technologically very important optical wavelength of 1.55μm. The use of the EPM for these calculations also leads to an unexpected effect in the response of the ground state energy to the applied electric field. The second half of the chapter is concerned with the 'stepped' asymmetric well formed

by sandwiching thin layers of InAs and GaSb between AlSb barriers. This system is particularly interesting not just because of the asymmetric quantum well formed in the conduction band but also because the InAs conduction and GaSb valence bands overlap.

Chapter 5

Chapter 5 covers the development of a method for extracting spatially slowly varying envelope functions, for quantum well bound states, from wavefunctions calculated with the EPM (using the complex band structure approach described in Chapter 3). These envelope functions are uniquely defined and obey the same (uncontroversial) interface boundary conditions as the original wavefunction. Extensive checks are performed to demonstate the validity of the theoretical and numerical procedures used in the envelope function extraction process.

Chapter 6

This chapter takes envelope functions produced by the method described in the previous chapter and discusses their properties, in particular, the behaviour of their spatial derivative around abrupt heterojunctions. For a particular bound state, direct comparison is made between the EPM derived envelope function of greatest magnitude and the (single) envelope function produced by a simple effective mass model.

Next, the dipole matrix elements between quantum well bound states are calculated employing the full EPM pseudo-wavefunction and various approximations to this, using the envelope function expansion with only certain terms included. The same approach is used to investigate interband charge oscillation, especially the effects of various envelope function approximations on the predicted results, as

compared to those predicted by the full EPM calculation.

Finally, a comparison is made between EPM valence band quantum well band structure and that calculated using so-called 'symmetrised' and 'ordered' $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians.

Summary and conclusion

Finally, we summarise the work presented in this thesis, recapping on the main conclusions, and suggesting future research areas that might lead on from this work.

Chapter 2

The empirical pseudopotential method

2.1 Introduction

The foundation for almost all the work presented in this thesis is the empirical pseudopotential method (EPM) for the calculation of the electronic structure of semiconductors. This is a powerful technique for the study of valence electrons and has been used extensively for both bulk and heterostructure calculations. Valence electrons are those that are in the outer shells of the atoms in the crystal.[1] At absolute zero, these electrons completely fill the valence band in a semiconductor, with the higher energy conduction bands empty. In contrast, electrons tightly bound to the atom are said to be in inner or core states. The aim of pseudopotential theory is to facilitate the solution of the Schrödinger equation for real materials, where one is interested in the valence rather than core electrons. This chapter outlines the basic theory and methodology, as applied to the current work, and presents results for bulk semiconductors which illustrate the important advantages and disadvantages

[1] First principles techniques often include lower energy electrons from closed shells but are not discussed here.

of this approach.

2.2 Theory

2.2.1 Brief review of pseudopotential theory

The term 'pseudopotential method' can be considered to cover two distinct approaches to the calculation of electronic structure, both originating from the same underlying theory. These are the empirical pseudopotential method (EPM) and *ab initio*, or first principles, methods. The current work uses the empirical pseudopotential method and so *ab initio* methods are not discussed here (see [23] for a review of such methods). The first step in the development of a pseudopotential theory is to make the Hartree, or mean-field, approximation[2] to reduce the many electron problem to that of solving the one-electron (time-independent) Schrödinger equation,

$$\widehat{H}\psi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m_e}\nabla^2 + V\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{2.1}$$

where m_e and E are the rest mass and energy of the electron, and V is the full self-consistent potential seen by the electron. The difficulty in solving (2.1) directly arises from this potential term. In a crystalline material, such as the group III-V and II-VI semiconductors considered in this work, the potential, V, has the symmetry properties of the crystal lattice. Since the mean-field potential depends partly on the electron-ion interactions, it varies rapidly over small distances around the ionic centres, but between the ions is relatively weak and slowly varying (we also note

[2] As the term *mean-field* suggests, this approximation assumes that each electron experiences an identical mean potential, V. The full wavefunction is then a Slater determinant of the one electron wavefunctions, solved using this mean potential, with the Pauli exclusion principle requiring each electron to be in a distinct quantum state. For further general explanation see, for example, p14 of [3] and chapter 12 of [1].

that because the potential includes a component due to the electrons themselves, a full, first principles solution to (2.1) requires a self-consistent approach).

For any real semiconductor, the form of V is such that no analytic solution to (2.1) is possible, and it is necessary to make a suitable approximation or use a numerical method, or some combination of the two. In choosing a numerical method, consideration must be given to the computational demands. For example, the periodicity of the crystal suggests the use of a standard Fourier series representation of the potential, V, and wavefunction, $\psi(\mathbf{r})$, as the basis of a numerical solution. However, due to the large and rapid variations in the potential the number of terms required to achieve even approximate convergence makes this approach unattractive.

An alternative approach which is more amenable to approximation than Fourier or plane wave methods is the tight-binding method (see, for example, Chapter 3 of [24]). This method approximates the crystal wavefunction by a linear combination of atomic orbitals (LCAO). This reproduces the valence band states satisfactorily, but fails to adequately represent excited states, i.e., the conduction bands, which are of critical importance in optical processes.

At the other extreme of approximations is the nearly free-electron model, or NFE (see, for example, p12 of [25]), to which the pseudopotential method is closely related. However, the NFE over-simplifies the situation, since it produces electronic structure that differs from the parabolic free-electron energy only in the immediate vicinity of the Brillouin zone boundaries. The pseudopotential approach bridges the gap between the LCAO and NFE pictures and has proven useful and reliable in many applications.[3] One further approach is the $\mathbf{k} \cdot \mathbf{p}$ method. This is a perturbation technique, which does not provide explicit knowledge of the periodic part of the Bloch function, but only the slowly varying envelope. Information about the periodic

[3]The pseudopotential method was derived from the orthogonalised plane wave (OPW) method, as first applied to the calculation of the electronic structure of solids by Herring [26].

part is included only through momentum matrix element parameters, which can be related to macroscopically determind effective masses. Pseudopotential theory in general, and the empirical pseudopotential method in particular, is discussed in detail by many authors in the literature (see, for example, [25,27,28]). Here we reproduce the key equations and highlight the important issues of the theory.

Pseudopotential theory requires the electrons to be split into two groups; *core* and *valence* electrons, as discussed in Section 2.1. In this context, we group the nucleus and core electrons together under the term *crystal ion*, and treat these ions as forming a fixed crystal lattice. The core electrons of one atom interact little with those of neighbouring atoms since they are tightly bound around the atomic nucleus in closed shells, thus, they can be well approximated by free atom states. The valence electrons are those in the open outer atomic shells of the (individual) atoms and these form the inter-atomic bonds. In this work, it is the valence electrons that are of primary interest, since as well as forming the covalent bonds, they are involved in charge transport and can undergo optical transitions.

The starting point in pseudopotential theory is the expression of the true wave-function, ψ, as a linear combination of the core states, ψ_c, and some smooth function, φ, to be determined,

$$\psi(\mathbf{r}) = \varphi(\mathbf{r}) + \sum_c b_c \psi_c(\mathbf{r}). \tag{2.2}$$

Here, we are only interested in valence states, so we require ψ to be orthogonal to the core states, i.e.,

$$\langle \psi | \psi_c \rangle = 0, \quad \text{for all } c. \tag{2.3}$$

Imposing this condition on ψ gives

$$\psi(\mathbf{r}) = \varphi(\mathbf{r}) - \sum_c \psi_c \langle \psi_c | \varphi \rangle. \tag{2.4}$$

On substitution into (2.1), the Schrödinger equation can be expressed in the form (see Appendix A)

$$\widehat{H}^p \varphi = \left(-\frac{\hbar^2}{2m_e} \nabla^2 + \widehat{W} \right) \varphi = E\varphi, \tag{2.5}$$

which defines the the pseudo-Hamiltonian, $\widehat{H}^p$, and the pseudopotential, $\widehat{W}$, which are given in terms of the core state projection operator, $\widehat{P}$, by

$$\widehat{W} = V + \sum_c (E - E_c) \widehat{P} \tag{2.6}$$

$$= V + \sum_c (E - E_c) |\psi_c\rangle\langle\psi_c|, \tag{2.7}$$

with the E_c being the energy eigenvalues of the core states. The function φ is called the pseudo-wavefunction. It can be shown that the two terms in the expression for the pseudopotential almost cancel [29,30], hence, although the true potential, V, varies greatly and rapidly with position, the pseudopotential is a relatively weak and slowly varying function. Thus a comparatively small number of terms may be used in a series expansion of the pseudo-wavefunction, (2.9), whilst still achieving reasonable convergence of the solution (see Section 2.6).

The pseudo-wavefunction still satisfies Bloch's theorem,

$$\varphi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{2.8}$$

where the Bloch periodic functions, $u_{n\mathbf{k}}(\mathbf{r})$, are periodic with the crystal lattice, and we have introduced the band index, n, and electron wave vector, $\mathbf{k}$, to distinguish

the states. Thus, it is natural to expand the pseudo-wavefunction as a Fourier series,

$$\varphi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{g}} a_{n\mathbf{g}\mathbf{k}} e^{i\mathbf{g}\cdot\mathbf{r}}, \tag{2.9}$$

with Fourier coefficients $a_{n\mathbf{g}\mathbf{k}}$, corresponding to the set of reciprocal lattice vectors $\{\mathbf{g}\}$.

A number of important points can now be made.

- The eigenvalue of $\widehat{H}^p$ is the same as that of $\widehat{H}$, i.e., the total energy of the electron, $E_{n\mathbf{k}}$ ($\mathbf{k}$ is included to indicate that the energy varies parametrically with the electron wave vector).

- The pseudo-Hamiltonian contains the kinetic energy operator, as before, but the potential term, V, has been replaced by the pseudopotential, $\widehat{W}$.

- The second term in $\widehat{W}$, involving the core projection operator has a repulsive effect on the valence states in and around the core region.

- The true wavefunction is related to the pseudo-wavefunction by

$$\psi_{\mathbf{k}}(\mathbf{r}) = \left(1 - \widehat{P}\right) \varphi_{\mathbf{k}}(\mathbf{r}).$$

  Therefore, in the regions between the atoms, where the valence electrons are concentrated, the pseudo-wavefunction is a good approximation to the true wavefunction of these states, since there is little overlap with the core states which are concentrated around the nuclei.

It is for the above reasons that the pseudopotential method is so powerful. However, solving (2.5) exactly is no more tractable than solving the original Schrödinger equation, since the core state energies and wavefunctions are required. For this reason empirical information is employed to further simplify the problem.

2.2.2 Empirical pseudopotentials

The empirical pseudopotential method (EPM) avoids the need for explicit knowledge of the core states by treating the pseudopotential as a parameterised function and adjusting the parameters until calculated band structure and experimental information on the actual band structure are in satisfactory agreement. It is clear that the pseudopotential, (2.7), is spatially non-local in general. It is possible to include non-local effects within the EPM scheme for bulk materials [31] but this is not feasible for heterostructure calculations if the complex band structure method is used (see Section 2.5), and so $\widehat{W}$ is replaced by a simple local pseudopotential, $V^p(\mathbf{r})$. This has the periodicity of the lattice and so can be expanded as a Fourier series using reciprocal lattice vectors, $\mathbf{g}$. So, we have

$$\widehat{W} \Rightarrow V^p(\mathbf{r}) = \sum_{\mathbf{g}} v_{\mathbf{g}} e^{i\mathbf{g}\cdot\mathbf{r}}, \tag{2.10}$$

where the $v_{\mathbf{g}}$ are Fourier coefficients, which may be expressed in terms of atomic pseudopotentials for the two atoms of the basis (in the case of diamond or zinc-blende materials).

We define the Bravais lattice points to be midway between the two atoms, labelled A and B, of the primitive unit cell of the zinc-blende crystal. Therefore, A and B are at $-\boldsymbol{\tau}$ and $+\boldsymbol{\tau}$ with respect to the Bravais lattice point, where $\boldsymbol{\tau} = (1,1,1)a/8$, and a is the lattice parameter. So, if $V_A^p(\mathbf{r})$ and $V_B^p(\mathbf{r})$ are the *atomic* pseudopotentials of A and B, the Fourier coefficients can be expressed as

$$v_{\mathbf{g}} = \frac{1}{\Omega_0}\left(e^{i\mathbf{g}\cdot\boldsymbol{\tau}}\int_{\Omega} V_A^p(\mathbf{r})e^{-i\mathbf{g}\cdot\mathbf{r}}d^3\mathbf{r} + e^{-i\mathbf{g}\cdot\boldsymbol{\tau}}\int_{\Omega} V_B^p(\mathbf{r})e^{-i\mathbf{g}\cdot\mathbf{r}}d^3\mathbf{r}\right), \tag{2.11}$$

where Ω_0 and Ω are the volumes of the primitive unit cell and crystal, respectively.

The integrals in (2.11) give the Fourier coefficients for the atomic pseudopotentials,

called the atomic form factors, $v_j(\mathbf{g})$, where $j = A$ or B. The atomic form factors can be defined for any point $\mathbf{q}$ in reciprocal space by the equation

$$v_j(\mathbf{q}) = \frac{1}{\Omega_j} \int_\Omega V_j^p(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d^3\mathbf{r}, \tag{2.12}$$

where $\mathbf{q}$ is a continuous wave vector and Ω_j is the 'atomic volume' of atom j in the crystal.[4] Here the atoms form a crystal lattice and the form factors are only required at certain discrete values of $\mathbf{q}$,

$$v_j(\mathbf{q}) \Rightarrow v_j(\mathbf{g}) = \frac{1}{\Omega_j} \int_\Omega V_j^p(\mathbf{r}) e^{-i\mathbf{g}\cdot\mathbf{r}} d^3\mathbf{r}. \tag{2.13}$$

It is common to assume that the atomic pseudopotentials are spherically symmetric, in which case it can be shown that the Fourier coefficient depends only on the radial component of the wave vector, i.e.,

$$v_j(\mathbf{q}) \rightarrow v_j(|\mathbf{q}|). \tag{2.14}$$

This greatly reduces the number of independent form factors that must be determined and so simplifies the fitting procedure (cf. Section 2.4.1). The general form of $v_j(|\mathbf{q}|)$ is shown schematically in Figure 2.1. Finally, it is convenient to express $v_\mathbf{g}$ in terms of symmetric and antisymmetric pseudopotential form factors, defined by

$$\begin{aligned} v^s_{|\mathbf{g}|} &= \frac{1}{2}\big(v_A(|\mathbf{g}|) + v_B(|\mathbf{g}|)\big) \\ v^a_{|\mathbf{g}|} &= \frac{1}{2}\big(v_A(|\mathbf{g}|) - v_B(|\mathbf{g}|)\big), \end{aligned} \tag{2.15}$$

[4]The 'atomic volume' is an ill-defined quantity that represents the volume of the primitive unit cell, Ω_0, occupied by each atom. Here the standard designation of $\Omega_A = \Omega_B = \Omega_0/2$ is used.

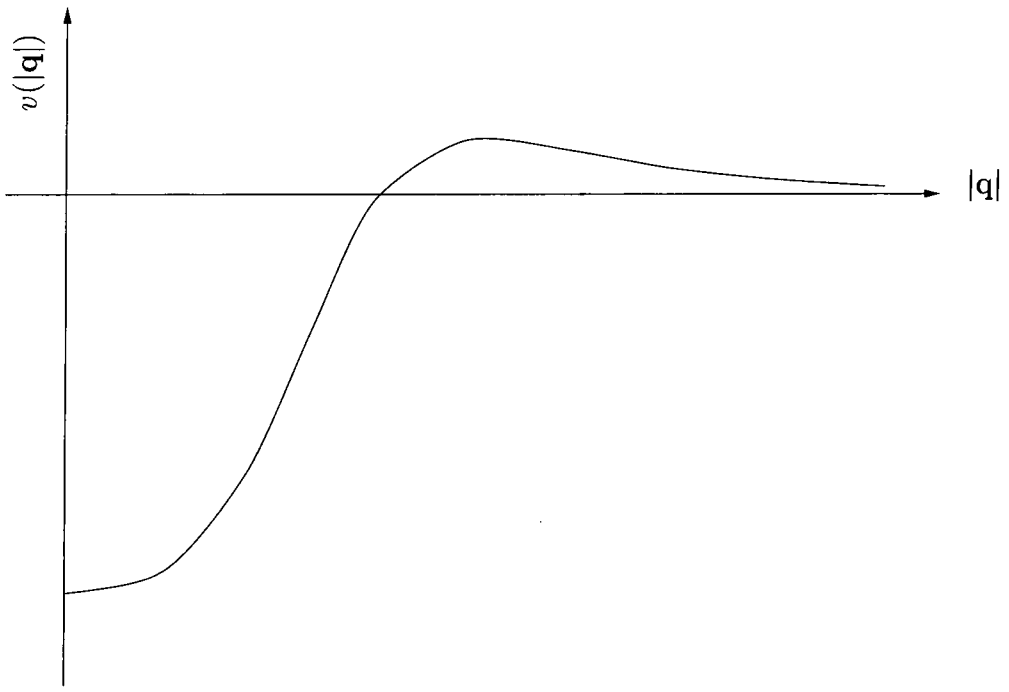


Figure 2.1 Schematic plot of the Fourier coefficient, $v(|\mathbf{q}|)$, for a typical atomic pseudopotential.

so that (2.11) becomes

$$v_{\mathbf{g}} = s^s_{\mathbf{g}} v^s_{|\mathbf{g}|} + i s^a_{\mathbf{g}} v^a_{|\mathbf{g}|}, \tag{2.16}$$

where s^s and s^a are called the symmetric and antisymmetric structure factors, and are given by

$$s^s_{\mathbf{g}} = \cos(\mathbf{g} \cdot \boldsymbol{\tau}) \quad \text{and} \quad s^a_{\mathbf{g}} = \sin(\mathbf{g} \cdot \boldsymbol{\tau}). \tag{2.17}$$

To summarise, the local EPM approximates the exact pseudopotential with $V^p(\mathbf{r})$, where

$$V^p(\mathbf{r}) = \sum_{\mathbf{g}} \left(s^s_{\mathbf{g}} v^s_{|\mathbf{g}|} + i s^a_{\mathbf{g}} v^a_{|\mathbf{g}|} \right) e^{i\mathbf{g}\cdot\mathbf{r}}. \tag{2.18}$$

Substituting the local pseudopotential into (2.5), and using the pseudo-wave-

functions, (2.9), as the basis, the pseudo-Hamiltonian has matrix elements

$$H_{\mathbf{gg'}} = \frac{\hbar^2}{2m_e}(\mathbf{k}+\mathbf{g})^2\delta_{\mathbf{gg'}} + v_{\mathbf{g}-\mathbf{g'}}. \tag{2.19}$$

The eigenvalues of $H_{\mathbf{gg'}}$ are the energies, $E_{n\mathbf{k}}$, with each corresponding eigenvector containing the pseudo-wavefunction Fourier coefficients, $a_{n\mathbf{gk}}$, for each reciprocal lattice vector. If n reciprocal lattice vectors are used in the pseudo-wavefunction expansion, there will be n eigensolutions, hence, to specify an energy band uniquely requires a band index, n. In this work, the band index is usually dropped unless its omission could lead to confusion.

Since the pseudopotential is weak and varies smoothly, its Fourier expansion may be truncated after only a small number of terms. In fact, only coefficients with $|\mathbf{g}| \leq \sqrt{11}(2\pi/a)$ are included, i.e., those for the reciprocal lattice vectors of type (000), {111}, {200}, {220} and {311} [32]. As the term with $\mathbf{g} = (000)$ merely contributes a spatially constant potential, v_0, it can be set to zero, or any other convenient value. As a result of this simplification, and the requirement that the structure factors be non-zero, the expansion includes only six parameters,

$$\begin{matrix} v_3^s & v_8^s & v_{11}^s \\ v_3^a & v_4^a & v_{11}^a, \end{matrix} \tag{2.20}$$

where we have used the more compact notation

$$v_{|\mathbf{g}|}^{s,a} \equiv v_{|\mathbf{g}|^2}^{s,a}. \tag{2.21}$$

The fitting of these to experiment is discussed in Section 2.4.1.

2.2.3 Spin orbit coupling

To include spin orbit coupling (SOC) effects, one modifies the pseudo-Hamiltonian by adding a term derived from an approximation of the Dirac equation (see, for example, p78 of [33]). This has been incorporated into the pseudopotential scheme by other workers [34–36], and here we use the simplest form of the correction, which has plane wave matrix elements given by

$$-i\lambda S_{\mathbf{g}-\mathbf{g}'}\,\boldsymbol{\sigma}\cdot(\mathbf{k}+\mathbf{g})\wedge(\mathbf{k}+\mathbf{g}') \tag{2.22}$$

where

$$\lambda S_{\mathbf{g}-\mathbf{g}'} = \left[-i\lambda_s\cos\big((\mathbf{g}-\mathbf{g}')\cdot\boldsymbol{\tau}\big)+\lambda_a\sin\big((\mathbf{g}-\mathbf{g}')\cdot\boldsymbol{\tau}\big)\right] \tag{2.23}$$

and $\boldsymbol{\sigma}$ is the Pauli spin matrix with Cartesian elements

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.24}$$

The parameters λ_s and λ_a are determined from experiment ($S_{\mathbf{g}}$ is the usual structure factor) and are usually expressed in terms of two further parameters,

$$\lambda_s = \frac{S_\mu(S_\alpha+1)}{2} \qquad \lambda_a = \frac{S_\mu(S_\alpha-1)}{2}. \tag{2.25}$$

S_μ determines the strength of the spin orbit coupling and S_α characterises the difference in strength of the spin orbit coupling in each atom of the primitive unit cell.

Thus, the pseudo-Hamiltonian now has matrix elements

$$\begin{aligned} H_{\mathbf{g}\mathbf{g}'ss'} = &-\frac{\hbar^2}{2m_e}(\mathbf{k}+\mathbf{g})^2\delta_{\mathbf{g}\mathbf{g}'}\delta_{ss'} + v_{\mathbf{g}-\mathbf{g}'}\delta_{ss'} \\ &+ \left[-i\lambda_s \cos\big((\mathbf{g}-\mathbf{g}')\cdot\boldsymbol{\tau}\big) + \lambda_a \sin\big((\mathbf{g}-\mathbf{g}')\cdot\boldsymbol{\tau}\big)\right] \\ &\times \boldsymbol{\sigma}_{ss'}\cdot(\mathbf{k}+\mathbf{g})\wedge(\mathbf{k}+\mathbf{g}'), \end{aligned} \tag{2.26}$$

where s and s' label the elements of the spin matrix. The eigenvectors of (2.5) are now spinors and the pseudo-wavefunction of (2.9) can be represented as

$$\varphi_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{g}} \begin{pmatrix} a_{\mathbf{g}\mathbf{k}\uparrow} \\ a_{\mathbf{g}\mathbf{k}\downarrow} \end{pmatrix} e^{i(\mathbf{k}+\mathbf{g})\cdot\mathbf{r}}. \tag{2.27}$$

As with the potential form factors, the spin parameter is adjusted to provide the best fit between the calculated and experimental results (see Section 2.4.2).

2.3 Inclusion of strain effects

2.3.1 Lattice mismatch in heterostructures

Strain effects are present in some of the heterostructures considered in this thesis because they consist of materials whose equilibrium lattice parameters are not equal. The lattice mismatch is illustrated schematically in Figure 2.2(a), where an epilayer, l, is to be grown on a substrate, s, whose lattice parameter is considered to be fixed at a_s.

Under such circumstances, the atoms of material l may align themselves with those of s in the plane of the interface. Thus, l forms a strained layer where there is biaxial stress in the plane of the interface. To compensate for the in-plane strain, l relaxes in the direction perpendicular to the interface, as shown in Figure 2.2(b). The

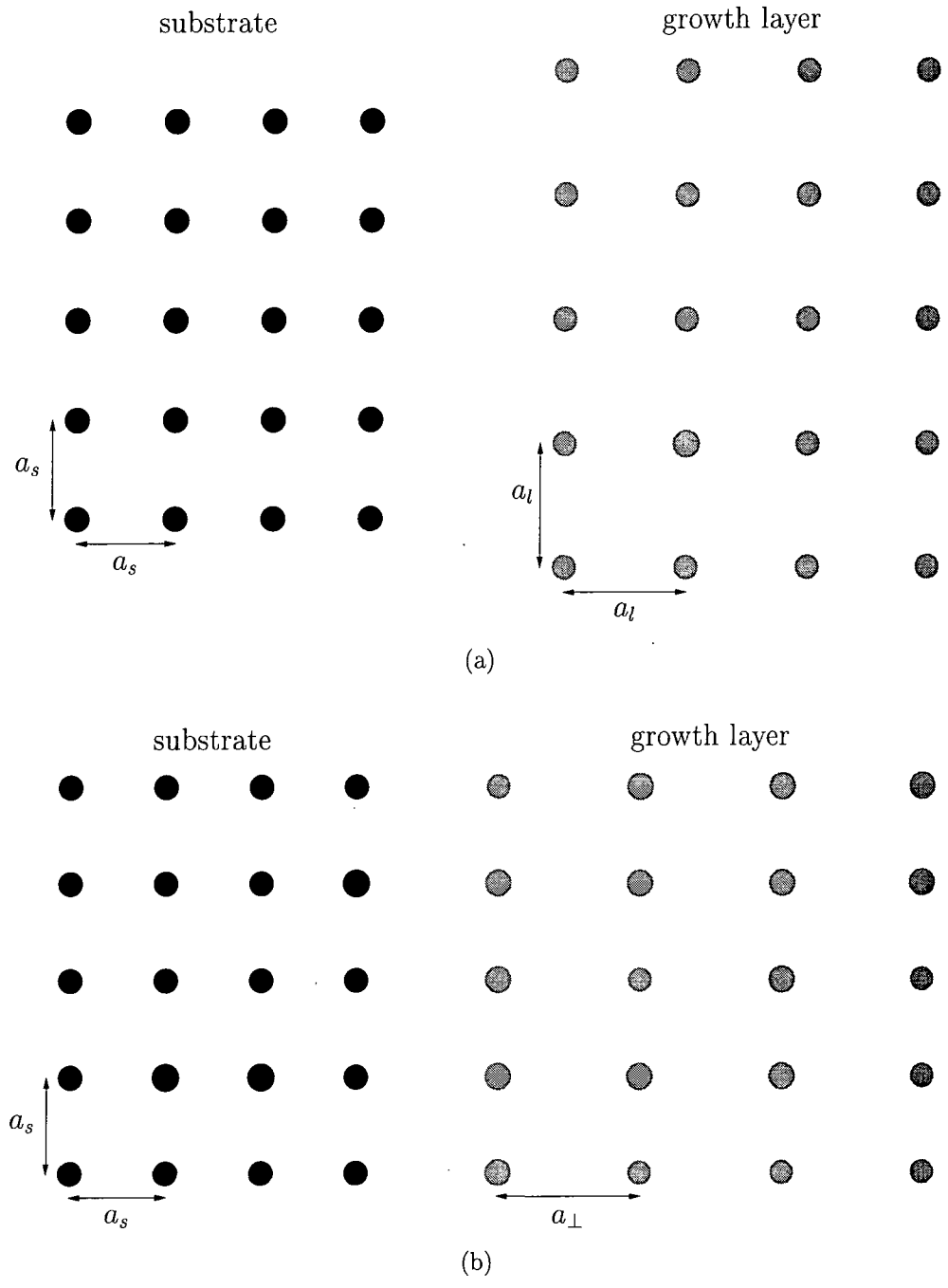


Figure 2.2 Schematic illustration of substrate and epilayer bulk lattice mismatch and strained layer growth. (a) The bulk equilibrium lattice parameters are not equal — in this example, the epilayer has the larger atomic spacing, $a_l > a_s$. (b) Dislocation free growth of the epilayer on the substrate. The epilayer relaxes in the growth direction, so that $a_\perp > a_l > a_s$.

cubic symmetry of the unit cells in l becomes tetragonal with two lattice parameters, $a_{||} = a_s$ and $a_\perp$. This situation, called the strained layer regime, is energetically favourable if the layer of l has a thickness less than the 'critical layer thickness'. For thicker layers we enter the relaxed regime and the appearance of dislocations is favoured. Here we consider only the strained layer regime, with perfect alignment of the atoms in the plane of all interfaces. First, the critical layer thickness is discussed in a little more detail.

Critical layer thickness

The critical layer thickness has been discussed in more detail by other workers elsewhere (see, for example, [37]). Here we present a brief review of the topic, with reference to the strained materials considered in this thesis.

Unfortunately, there is still not complete agreement on the way that the critical layer thickness, h_c, depends on the material properties and the in-plane strain. One prediction is [38]

$$h_c = \left(\frac{1-\nu/4}{4\pi(1+\nu)}\right) b\epsilon_{||}^{-1}\big(\ln(h_c/b)+\theta\big), \tag{2.28}$$

where ν is Poisson's ratio (typically $\approx \frac{1}{3}$), b is the dislocation Burger vector (≈ 4Å), $\epsilon_{||}$ is the in-plane strain and $\theta \approx 1$. The magnitudes of the in-plane strains for InAs and AlSb matched to GaSb are both approximately 0.65%, which gives $h_c \approx$ 158Å, or $\epsilon_{||}h_c \approx 100$Å%. Experimental investigations of $In_xGa_{1-x}As$ layers on GaAs have resulted in $\epsilon_{||}h_c$ values of less than 100Å% [39] and more recently 200Å% [40], suggesting that (2.28) is not unreasonable. However, one should note that h_c also depends on the growth temperature and rate [41,42].

2.3.2 Elastic theory and heterostructures

The electronic energy band structure of a semiconductor may be significantly altered if it is stressed. The stress causes the material to be strained, i.e., the atoms of the crystal lattice structure are displaced with respect to their equilibrium positions. Hydrostatic stress does not affect the crystal symmetry but merely changes the lattice parameter. However, any other type of stress will lower the cubic symmetry of the (zinc-blende) crystal and therefore lower the symmetry of the energy bands in reciprocal space.

The distortion of the stressed material is given by the strain tensor, ϵ_{ij}. For sufficiently small stress the elastic response is linear and the strain tensor is then related to the stress tensor by

$$\epsilon_{ij} = \sum_{kl} s_{ijkl}\, \sigma_{kl}, \tag{2.29}$$

where s_{ijkl} is a fourth-rank tensor called the elastic compliance constant. It can be shown (if there is no net torque) that both the stress and strain tensors are symmetric, in which case it is convenient to adopt a matrix notation (see, for example, [43]). This expresses the stress and strain tensors as column matrices, σ_i and ϵ_i, containing the six independent elements of their respective tensors. The fourth-rank compliance tensor can then be represented by a square symmetric matrix, S, of order six, with elements s_{ij}, i.e.,

$$\epsilon_i = \sum_{j} s_{ij}\, \sigma_j. \tag{2.30}$$

The elements s_{ij} are referred to as the compliance constants. In cubic crystals (including diamond and zinc-blende), symmetry considerations show that there are in fact only three independent compliance constants, labelled s_{11}, s_{12} and s_{44}, with

the remaining non-zero elements given by [43]

$$\begin{aligned} s_{11} &= s_{22} = s_{33} \\ s_{12} &= s_{13} = s_{23} \\ &\left(= s_{21} = s_{31} = s_{32} \right) \\ s_{44} &= s_{55} = s_{66}. \end{aligned} \tag{2.31}$$

The inverse of the matrix S contains the elastic moduli (or stiffness constants), c_{ij}, i.e.,

$$\sigma_i = c_{ij}\,\epsilon_j, \tag{2.32}$$

whose three independent elements are related to the compliance constants by [43]

$$c_{11} = \frac{s_{11} + s_{12}}{(s_{11} - s_{12})(s_{11} + 2s_{12})} \qquad c_{12} = -\frac{s_{12}}{(s_{11} - s_{12})(s_{11} + 2s_{12})} \qquad c_{44} = \frac{1}{s_{44}}. \tag{2.33}$$

The compliance constants and elastic moduli are macroscopic properties of each bulk material, and experimentally determined values are readily available [44].

Three distinct types of stress (and hence strain) are considered — hydrostatic, uniaxial and biaxial. The hydrostatic and uniaxial stresses are used in the fitting procedure (see Sections 2.3.3 and 2.4.3), whilst the biaxial stress is the type actually present in the lattice mismatched heterostructures to be modelled. These stresses are further limited to be acting along the crystal axes, i.e., perpendicular to the faces of the fcc conventional unit cell. In all three cases, the shear strain elements, ϵ_4, ϵ_5 and ϵ_6, are zero, which has the consequence that the faces of the conventional unit cell remain mutually perpendicular [5]. This is not necessary for the application

[5] In fact, the critical restriction on the stress is that there are stresses acting along the con-

of the EPM used here, but is helpful since the components of the reciprocal lattice vectors after stress is applied, g_i^{s}, are simply related to the unstressed case, g_i^{u}, by

$$g_i^{\mathrm{s}} = g_i^{\mathrm{u}} \frac{a_i^{\mathrm{u}}}{a_i^{\mathrm{s}}} \qquad i = x, y, z, \tag{2.34}$$

where a_i^{s} and a_i^{u} are the strained and unstrained lattice parameters in the i^{th} direction. Since the heterostructures under consideration in this work are all grown in the $\langle 100 \rangle$-type direction this restriction on the stresses does not pose a problem.

Hydrostatic stress

For the hydrostatic case, $\sigma_{ij} = \sigma\delta_{ij}$ and so

$$\epsilon_i = \begin{cases} (s_{11} + 2s_{12})\sigma & i = 1, 2, 3 \\ 0 & \text{otherwise.} \end{cases} \tag{2.35}$$

Uniaxial $\langle 100 \rangle$ stress

For uniaxial stress along the x-axis, the only non-zero stress tensor element is $\sigma_{11} = \sigma$, and so

$$\epsilon_i = \begin{cases} s_{11}\sigma & i = 1 \\ s_{12}\sigma & i = 2, 3 \\ 0 & \text{otherwise.} \end{cases} \tag{2.36}$$

ventional unit cell axis only. This allows a general triaxial stress, where $\sigma_{11} \neq \sigma_{22} \neq \sigma_{33}$, to be considered in the same way.

Uniaxial stress is used to define Poisson's ratio, ν, which is minus the ratio of the strains perpendicular and parallel to the stress axis,

$$\nu = -\frac{\epsilon_{2,3}}{\epsilon_1} = -\frac{s_{12}}{s_{11}} = \frac{c_{12}}{c_{11}+c_{12}}. \tag{2.37}$$

where ν is a property of the bulk crystal and can be determined experimentally for each material.

Biaxial stress

Since we are interested in biaxial stresses for heterostructure modelling, it is assumed that the stresses in both in-plane directions are equal, in which case the stress tensor takes the simple form

$$\sigma_{ij} \equiv \begin{bmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{2.38}$$

if the biaxial stress is in the xy-plane. In the matrix notation,

$$\sigma_i = \begin{cases} \sigma & i = 1,2 \\ 0 & \text{otherwise} \end{cases} \tag{2.39}$$

Thus, the strain is given by

$$\epsilon_i = \begin{cases} (s_{11}+s_{12})\sigma & i = 1,2 \\ 2s_{12}\sigma & i = 3 \\ 0 & \text{otherwise.} \end{cases} \tag{2.40}$$

So, in the notation used for the heterostructure description (see Section 2.3.1), the strains parallel and perpendicular to the growth plane are

$$\begin{aligned} \epsilon_{\parallel} &= (s_{11} + s_{12})\sigma \\ \epsilon_{\perp} &= 2s_{12}\sigma. \end{aligned} \tag{2.41}$$

Therefore, since in a heterostructure $\epsilon_{\parallel}$ is known, one can calculate $\epsilon_{\perp}$ directly from the ratio between the two strains. So, from (2.41) and (2.37), one obtains

$$\epsilon_{\perp} = -\frac{2\nu}{1-\nu}\epsilon_{\parallel}. \tag{2.42}$$

For the materials considered, Poisson's ratio typically has a value of about $\frac{1}{3}$, so that $\epsilon_{\perp} \approx -\epsilon_{\parallel}$.

2.3.3 Strain within the EPM

The effects of strain are introduced into the EPM chiefly via alterations to the pseudopotential term, $V^p(\mathbf{r})$, since this contains the information about the crystal structure. The kinetic energy and spin orbit coupling terms are only altered through the dependence of the wave and reciprocal lattice vectors on the strained lattice dimensions. From (2.11) one can see that the strain alters the pseudopotential terms in two ways. Firstly, $v_{\mathbf{g}}$ depends inversely on the primitive unit cell volume, which will in general be changed by the strain. Secondly, the components of the reciprocal lattice vectors are modified due to the change in the direct lattice vectors. Note that although both $\boldsymbol{\tau}$ and $\mathbf{g}$ are altered, their scalar product $\mathbf{g} \cdot \boldsymbol{\tau}$ is constant [45]. The pseudopotentials V_A^p and V_B^p of (2.11) are those of isolated atoms, so they are not affected by the introduction of strain and have a continuous representation in reciprocal space, as indicated in Figure 2.1 and by (2.12). Therefore, if the values of $\mathbf{g}$ are changed by straining the lattice, we merely require $v_j(|\mathbf{q}|)$ at different discrete

values of $|\mathbf{q}|$. The new values for the form factors are then scaled according to the change in the primitive unit cell volume, $\Omega_0 = 2\Omega_{A,B}$.

In practice, the symmetric and antisymmetric form factors are adjusted directly, since they can be thought of as values of continuous functions at certain points in Fourier space for an isolated pair of atoms, A and B. By convention, A and B are the cation and anion, respectively. So, for example, InP is modelled with indium as atom A and phosphorus as atom B. With this convention, the continuous symmetric and antisymmetric form factors have the form depicted in Figure 2.3. Note that at $|\mathbf{q}| = 0$, the symmetric form factor can be given, in rydbergs, by [27]

$$v^s(0) = -\frac{2}{3}E_F = -\frac{2}{3}k_F^2, \tag{2.43}$$

where

$$k_F = \left(\frac{3\pi^2 Z}{\Omega_0}\right)^{\frac{1}{3}}, \tag{2.44}$$

with Z being the average valency *per atom*, i.e., 4 for this work, and Ω_0 the primitive unit cell volume, $a^3/8$, in atomic units. So, to obtain the new form factor values for a given strain, we must merely determine the new values of the reciprocal lattice vectors at which v^s and v^a are required (remembering to scale the resultant form factor values according to the change in the unit cell volume). Unfortunately, this simplicity is spoiled by the fact that the crystal form factors are treated as adjustable parameters and we have no analytic expression for the continuous Fourier coefficients.

One way around this problem is to fit analytic curves that pass through each set of form factors [46–48]. This is appealing since the general form of the curves is known and it is straightforward to decide whether a fit is physically reasonable.

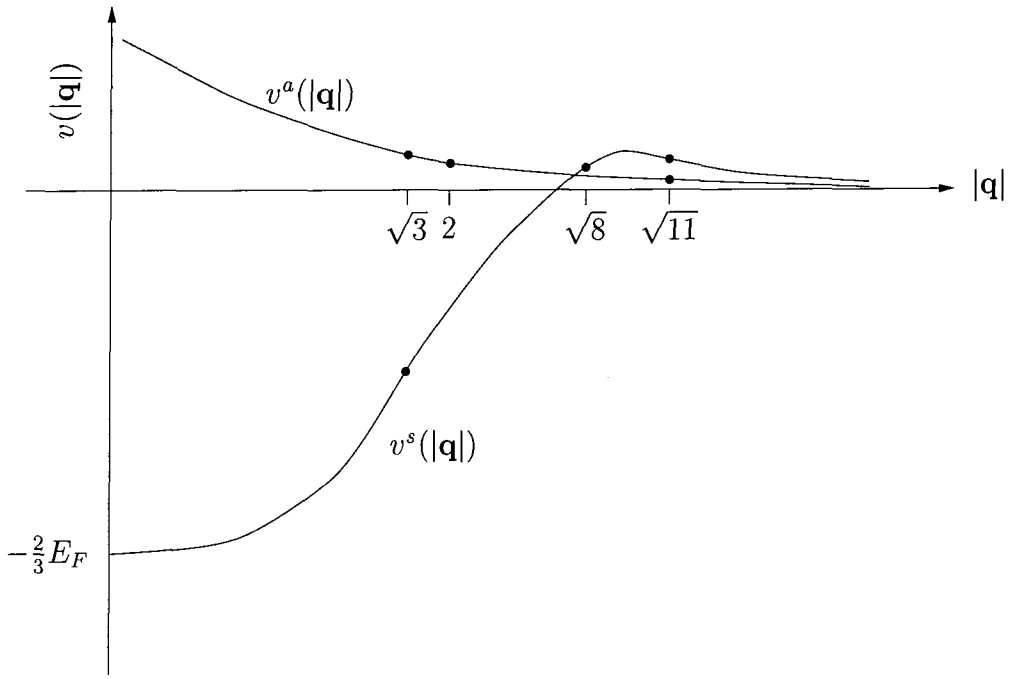


Figure 2.3 Schematic plot of typical symmetric and antisymmetric pseudopotential Fourier coefficients. The positions of each $|\mathbf{g}|$ are indicated, in units of $2\pi/a$. See text for details.

It also allows any amount of strain to be modelled, since the fitted curves will have values over all $|\mathbf{q}|$-space. However, whilst it might be possible to find satisfactory looking curves that pass through the values of $v^s_{|\mathbf{g}|}$ and $v^a_{|\mathbf{g}|}$, this does not guarantee the energy band structure will vary correctly with strain. For this reason, a second method is used where the form factor curves are assumed to be linear in the region of each reciprocal lattice vector [45,49]. The strain induces a change in the magnitude of each reciprocal lattice vector, $\Delta|\mathbf{g}|$, and so the form factors after straining are given by

$$v^{s,a}_{|\mathbf{g}|} \rightarrow \left(v^{s,a}_{|\mathbf{g}|} + g^{s,a}_{|\mathbf{g}|}\Delta|\mathbf{g}|\right)\frac{\Omega_0}{\Omega_0'}, \tag{2.45}$$

where $g^j_{|\mathbf{g}|}$ is the gradient of the Fourier coefficient passing through $v^j_{|\mathbf{g}|}$, and Ω_0 and Ω_0' are the primitive unit cell volumes before and after straining.

The gradients of these linear sections are adjusted so that the calculated band

structure strain dependence is in agreement with experiment, although extra conditions are applied so that the gradients are in line with the known form of the curves (see Section 2.4.3). This can only be a good approximation if the magnitudes of the reciprocal lattice vectors after straining are close to those of the unstrained crystal, i.e., if the strain is small[6], however, the systems studied in the present work are either equilibrium lattice matched or have only small strains (typically 1% or less), hence the linear approximation is expected to be reasonable.

2.4 Parameter fitting

Having discussed the theory behind the EPM, we turn to the procedures involved in fitting the various parameters. For all semiconductors, the energy band structure is dependent on temperature, the most important effect being the appreciable decrease in the band gap with increasing temperature (see, for example, [50]). Therefore it is necessary to decide at which temperature we wish to model the materials, and fit to the appropriate experimental data. Since much of the present work is aimed at heterostructures with possible device applications, all the materials considered are modelled at room temperature ($\sim$ 300K). The parameters that require fitting to experimental data are summarised in Table 2.1.

2.4.1 Form factors

The first parameters to be fitted are the pseudopotential form factors, which determine the energy band structure in the absence of spin orbit coupling. We wish to find values for the form factors that produce band structure that is in agreement with

[6]Strictly speaking, this is not necessarily the case. For example, a large uniaxial stress will lead to a large strain along the stress axis, but also strains of the opposite sign in the plane perpendicular to that axis, hence the magnitude of the reciprocal lattice vectors will not be greatly altered. It is only when the crystal is not allowed to relax elastically, such as when hydrostatic pressure is applied, that large strains lead to large changes in $|\mathbf{g}|$.

Form factors	v_3^s, v_8^s, v_{11}^s, v_3^a, v_4^a, v_{11}^a
Spin parameter	S_μ
Form factor gradients	g_3^s, g_8^s, g_{11}^s, g_3^a, g_4^a, g_{11}^a

Table 2.1 Adjustable parameters used for the EPM. The form factors determine the energy band structure without spin, the spin parameter specifies the strength of the spin orbit coupling, and the form factor gradients determine the behaviour of the band structure under strain.

experiment. In principle, one could compare the calculated and experimental results for a number of types of experimental data (such as density of states, reflectivity curves, etc), but in this work we have confined the comparison to energy transitions at points in the Brillouin zone of high symmetry. This is because the fitting is carried out without the inclusion of spin orbit coupling (SOC) in order to decrease the computational effort required. It is a straightforward task to 'spin-average' experimental transition energies at specific points in the Brillouin zone but not quantities that involve summations over states in the zone. Spin-averaging determines what the energies between bands would be in the absence of SOC. For example, in the absence of SOC, the top of the valence band (at Γ) is triply degenerate, with each energy band being doubly spin degenerate. If the SOC is finite the degeneracy is split into the light and heavy hole doublet and the spin split off (SSO) band (each of which is still doubly spin degenerate at Γ). This splitting is called the spin split off energy, or Δ_0. The light and heavy hole bands are higher in energy than the SSO band, and move up by $\Delta_0/3$ whilst the SSO drops by $2\Delta_0/3$ (see, for example, p15 of [50]). Hence, to spin-average the direct energy gap, the experimentally determined gap is increased by $\Delta_0/3$. At other positions in the Brillouin zone, such as the L and X points, the degeneracies are different and so the spin-averaging involves adding a different fraction of the degeneracy splitting at these points.

The fitting procedure uses a Monte Carlo technique, since the transition energies depend in some unknown way on the *six* form factors. This is described below, and

presented diagramatically in Figure 2.4.

First, the form factors are initialised and the transition energies at the required k-points are calculated. These are compared to the desired (spin-averaged) experimental energies and the magnitude of the errors is summed with a weighting for each transition. The weighting is used because it is generally not possible to reproduce all the transition energies accurately using a local EPM [31,51]. Those transitions that are considered more important than others, typically those involving the direct and indirect energy gaps, are given a higher weighting. Next, one of the six form factors is both selected and altered at random until the new set of form factors satisfy certain criteria (see below), then the transition energies are recalculated. The new total weighted error is determined and if it is less than the previous one, the new set of form factors is retained. This process is repeated until satisfactory results are obtained.

Criteria for valid form factors

The general form in reciprocal space of the pseudopotentials is known, see Figure 2.3, hence certain conditions can be imposed on the form factors. The value of v_3^s must be negative, with a larger magnitude than the other two symmetric form factors, since the potential decays with distance from the atom. For the group IV elemental crystals, the two atoms of the primitive unit cell are identical and hence the antisymmetric form factors are zero. For polar materials, with the atoms arranged as described in Section 2.3.3, v_3^a must be positive with the remaining v^a smaller in magnitude, again due to the decay of the atomic pseudopotentials. In principle, v_4^a and v_{11}^a could be negative with the latter even having the larger magnitude of the two, however, to reduce the size of the form factor phase space, we further restrict

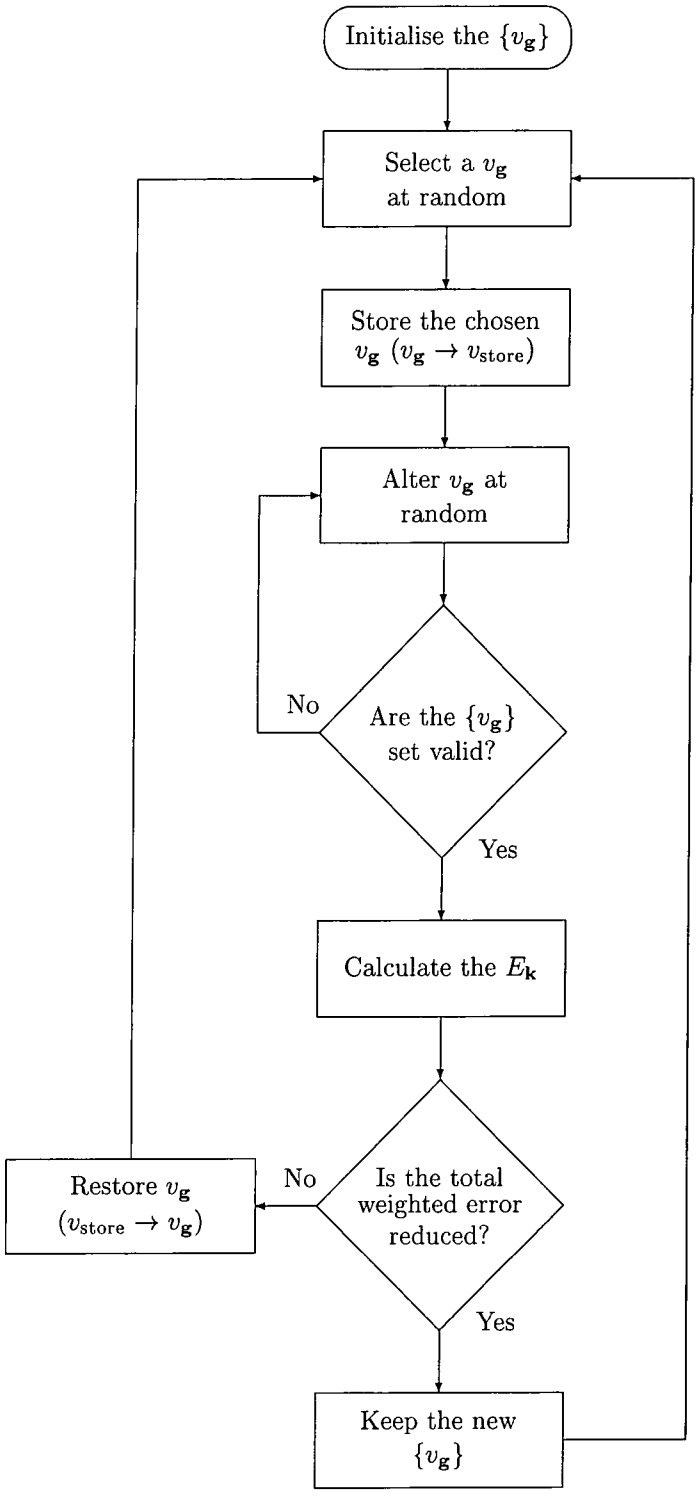


Figure 2.4 The empirical pseudopotential form factor fitting procedure. The set of six form factors is represented by $\{v_{\mathbf{g}}\}$. The process is repeated until the transition energies, $E_{\mathbf{k}}$, are in satisfactory agreement with experiment. See text for details.

these antisymmetric form factors to obey the commonly applied condition

$$v_3^a > v_4^a > v_{11}^a \geq 0. \tag{2.46}$$

2.4.2 Spin parameter

Having obtained satisfactory agreement with the spin-averaged experimental data, the spin orbit coupling is introduced. This is a straight forward procedure, as the spin split off energy, Δ_0, varies smoothly and monotonically with the spin parameter S_μ. Therefore, S_μ is adjusted so that the calculated and experimental values of Δ_0 are in agreement. When this has been done, the transition energies, in particular the direct gap, should have the correct size. In practice, the $+\Delta_0/3$ and $-2\Delta_0/3$ shifting of the light/heavy and SSO bands is not exact, so the spin-averaged energies used in the fitting may need to be adjusted slightly so that both the direct gap and the spin split off energy are correct.

2.4.3 Form factor gradients

The final fitting procedure is that of the form factor gradients. The linear approximation (see Section 2.3.3) means that there is one gradient per form factor, i.e., we have six further parameters (for polar materials). These are adjusted in the same way as the form factors themselves, except the fits are to experimental deformation potentials rather than transition energies, and spin *is* included explicitly in the fitting procedure. As with the form factor fitting, certain constraints can be applied to the gradients. Clearly, given the condition (2.46), the gradients for the antisymmetric form factors must satisfy

$$g_3^a < g_4^a < g_{11}^a \leq 0. \tag{2.47}$$

Also, g_3^s must be positive. However, v_8^s and v_{11}^s are near a turning point, so their gradients may be positive, negative or even zero.

2.5 Complex band structure

When calculating band structure, it is usual to seek all the allowed energies, $E_{n\mathbf{k}}$, for a given real electron wave vector, $\mathbf{k}$. However, one may also pick a (real) energy and find all the allowed wave vectors, which will in general have complex components. The resulting band structure is said to be *complex* because the energy is now a function of complex wave vector.[7] The heterostructure calculations presented in this thesis have been carried out using the complex band structure approach. These complex bands are not allowed in bulk materials, where the application of Born-von Karman (i.e., periodic) boundary conditions require the wave vector to be real (see, for example, p135 of [52]).

However, we may still write the wavefunction in the Bloch form,

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}), \tag{2.48}$$

where $u_{\mathbf{k}}(\mathbf{r})$ has the same periodicity as the lattice and $\mathbf{k}$ is in general complex.[8] If $\mathbf{k}$ is real, the probability density is $|u_{\mathbf{k}}(\mathbf{r})|^2$, which is clearly also periodic with the lattice, i.e., the exponential acts merely as a phase factor. However, suppose $\mathbf{k}$ has a non-zero imaginary z-component, $k = k_r + ik_i$, then the probability density is $e^{-2k_i z}|u_{\mathbf{k}}(\mathbf{r})|^2$, which diverges for $k_i z < 0$, thus violating the quantum mechanical boundary condition requiring the probability density to remain finite over all

[7] In general, the eigenstates of the pseudo-Hamiltonian will be complex, even if the wave vector is real. Hence, when a state is refered to as being complex, we mean that its wave vector has at least one non-real component.

[8] Alternatively, a heterostructure wavefunction can always be expressed as a linear combination of Bloch states with real $\mathbf{k}$ (see Chapter 5).

space. However, if a single interface is present the probability may decay *away* from this in either direction, so that complex-k states are valid. If there is more than one interface, decaying and growing states are valid within the region between the interfaces.

In this thesis, we are interested in 'two dimensional' heterostructures, i.e., those that confine the electrons in one direction. We consider structures grown along the $\langle 001 \rangle$ direction, so that only a single component of the wave vector may be complex. The growth direction is termed *perpendicular* ($\perp$), because it is perpendicular to the growth layers. The growth plane is called the *in-plane* or *parallel* ($\|$) direction. So, the wave vector can be expressed as

$$\begin{aligned} \mathbf{k} &= \mathbf{k}_{\|} + k\mathbf{e}_{\perp} \\ &= \mathbf{k}_{\|} + (k_r + ik_i)\mathbf{e}_{\perp}, \end{aligned} \tag{2.49}$$

where $\mathbf{e}_{\perp}$ is a unit vector in the growth direction. Since only bound states are investigated, the electron energy, $E_{\mathbf{k}}$, is restricted to real values, while the bulk nature of the in-plane direction requires that the wave vector components in the plane are real. Thus, we have

$$k \in \mathbb{C} \qquad \text{and} \qquad \mathbf{k}_{\|}, E_{\mathbf{k}} \in \mathbb{R}, \tag{2.50}$$

and rather than trying to find real energies for complex k and real $\mathbf{k}_{\|}$, it is more efficient to solve for those k that correspond to a given real $\mathbf{k}_{\|}$ and energy. To achieve this, the matrix elements of the Schrödinger equation, (2.26), are rearranged into terms which are independent of k and those dependent on k and k^2. This can then be rewritten as a matrix polynomial in k,

$$\left(-H^{(0)} - kH^{(1)}\right)\varphi = k^2\varphi, \tag{2.51}$$

where

$$H^{(0)}_{\mathbf{g}\mathbf{g}'ss'} = \left((\mathbf{k}+\mathbf{g})^2 - \frac{2m_eE}{\hbar^2}\right)\delta_{\mathbf{g}\mathbf{g}'}\delta_{ss'} + \frac{2m_e}{\hbar^2}v_{\mathbf{g}-\mathbf{g}'}\delta_{ss'} - \frac{2m_e}{\hbar^2}i\lambda S_{\mathbf{g}-\mathbf{g}'}\,\boldsymbol{\sigma}_{ss'}\cdot(\mathbf{k}_{||}+\mathbf{g})\wedge(\mathbf{k}_{||}+\mathbf{g}') \tag{2.52}$$

$$H^{(1)}_{\mathbf{g}\mathbf{g}'ss'} = 2g\delta_{\mathbf{g}\mathbf{g}'}\delta_{ss'} - \frac{2m_e}{\hbar^2}i\lambda S_{\mathbf{g}-\mathbf{g}'}\,\boldsymbol{\sigma}_{ss'}\cdot\mathbf{e}_\perp\wedge(\mathbf{g}'-\mathbf{g})$$

This can then be expressed as an eigenvalue equation, with the complex wave vector, k, as the eigenvalue [21,53],

$$\begin{pmatrix} 0 & \mathrm{I} \\ -H^{(0)} & H^{(1)} \end{pmatrix}\begin{pmatrix} \varphi \\ k\varphi \end{pmatrix} = k\begin{pmatrix} \varphi \\ k\varphi \end{pmatrix}, \tag{2.53}$$

where I is the identity matrix and we have made use of the trivial expression

$$0\times\varphi + k\times\varphi = k\times\varphi \tag{2.54}$$

to form the top row of the matrix.

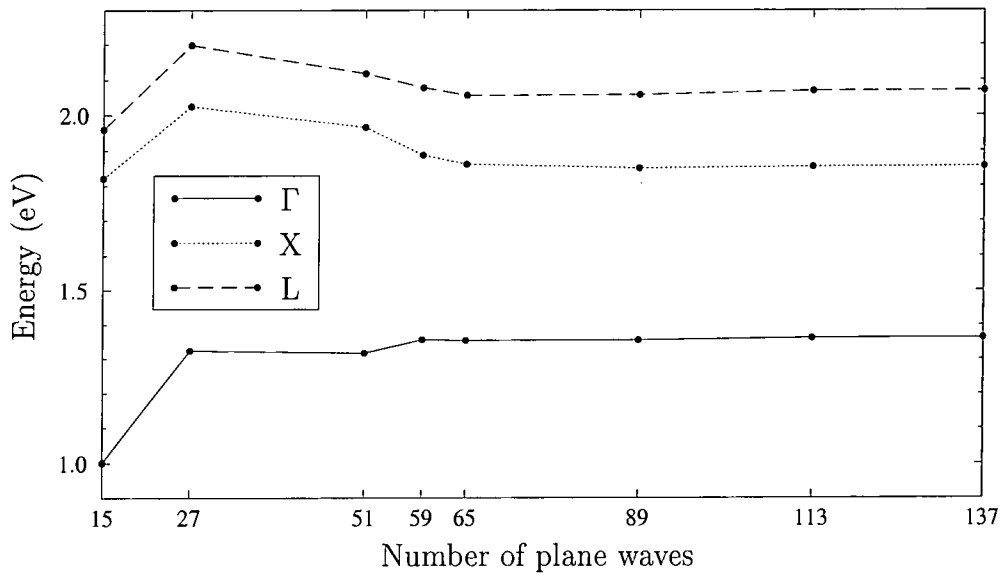
Any numerical calculations require the Fourier expansion of the pseudo-wavefunction, (2.9), to have a finite number of terms, therefore, the number of k eigenvalues in (2.53) will also be finite. With spin included, the order of the matrix in the eigenequation is four times the number of reciprocal lattice vectors used in the Fourier expansion, $n_{\mathbf{g}}$ (which is 89 for all cases in this work). Thus, there will be $4\times n_{\mathbf{g}}$ eigenvalues for k. However, some of these represent equivalent solutions, i.e., there are values of k separated by a reciprocal lattice vector. In fact, the number of independent solutions is given by $4\times n_{\mathbf{g}_{||}}$, where $n_{\mathbf{g}_{||}}$ is the number of different in-plane components, $\mathbf{g}_{||}$, of the $n_{\mathbf{g}}$ reciprocal lattice vectors, and spin is included [54,55].

2.6 Convergence

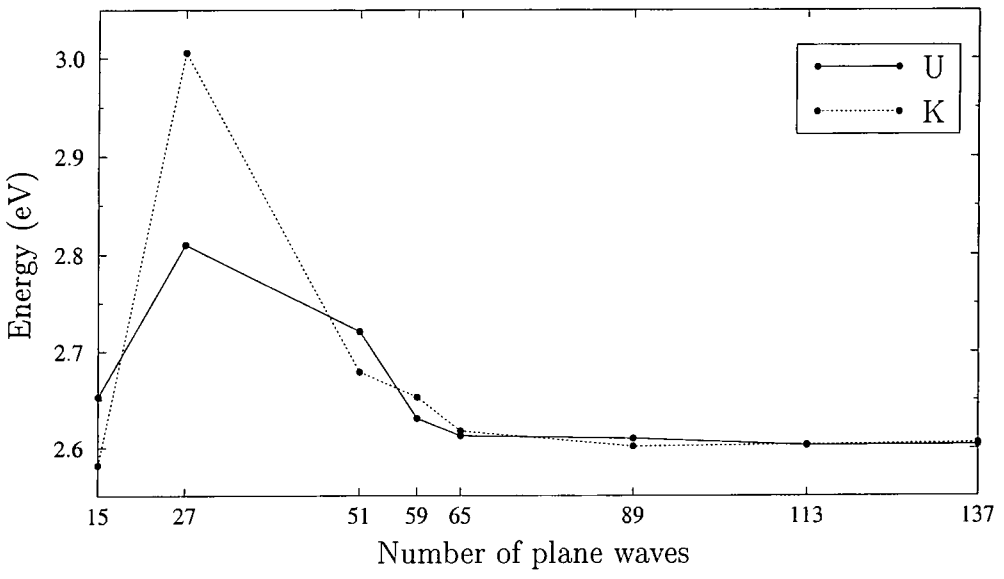
It was stated in Section 2.2.1 that the pseudo-wavefunction requires only a 'small' number of terms in its Fourier expansion for convergence. In this section, we discuss suitable values for this 'small' number.

The number of calculations necessary to solve the Hamiltonian matrix eigenequation, (2.19), and therefore the time taken by the digital computer to do so, is proportional to n^3, where n is the order of the matrix, i.e., the number of reciprocal lattice vectors, $\mathbf{g}$, used in the Fourier expansion of the pseudo-wavefunction, (2.9). Thus, using 137 rather than 89 $\mathbf{g}$'s results in calculations that take about $(\frac{137}{89})^3 \approx 3.6$ times longer. Since complex band structure calculations require the solution of eigenvalue equations of order four times the number of $\mathbf{g}$'s, it is clearly highly desirable to use as few $\mathbf{g}$'s as possible, whilst still achieving convergence adequate for the present purposes.

The variation in energy of the first conduction band with the number of plane waves used in the pseudo-wavefunction expansion, (2.9), is plotted in Figure 2.5. The energies are those at high symmetry points in the first Brillouin zone measured from the top of the valence band. Symmetry considerations require that the energies at U and K, $(1\frac{1}{4}\frac{1}{4})\frac{2\pi}{a}$ and $(\frac{3}{4}\frac{3}{4}0)\frac{2\pi}{a}$, respectively, should be equal. However, Figure 2.5(b) shows that if less than 65 plane waves are used, not only have the U and K energies not converged but they are quite different from each other. Even when 89 plane waves are used, the second conduction band displays noticeable dissagreement between the calculated U and K energies, as can be seen in Figure 2.6. However, due to the cubic dependence of the calculation time on the number of $\mathbf{g}$'s, a figure of 89 is considered to provide adequate convergence for this work.



(a)



(b)

Figure 2.5 Convergence of the first conduction band energy with increasing number of plane waves.

2.7 Examples of bulk calculations

This section contains results for various properties of indium phosphide (which forms the barriers for one of the quantum wells discussed in Chapter 6), calculated using the EPM as described in this chapter. The aim of this section is to demonstrate the accuracy of the local EPM. The calculations are for room temperature (which affects the sizes of the fundamental band gap and lattice parameter), using 89 plane waves (per spin component) in the wavefunction expansion, and are based on the parameters listed in Appendix D.

Real band structure

Figure 2.6 shows the electronic energy band structure for real wave vectors along lines of high symmetry in the first Brillouin zone. The band structure exhibits features typical of the zinc-blende materials, with energy band gaps between the lowest valence and the spin split off bands as well as between the valence and conduction bands. The top of the valence band is always at the zone centre, Γ. The lowest point of the conduction band may be either at Γ (*direct* gap material), as in this case, or some other point in the Brillouin zone (*indirect* gap material), as with AlSb, for example.

Density of states

The electron density of states per primitive unit cell at energy E can be expressed as (see, for example, p143 of [52])

$$g(E) = \sum_n g_n(E) = \sum_n \int \frac{d^3\mathbf{k}}{(2\pi)^3} \delta(E - E_{n\mathbf{k}}), \tag{2.55}$$

where the sum is over all energy bands[9], n, and the integral is over a single Brillouin zone. The density of states for indium phosphide is plotted in Figure 2.7, with the van Hove singularities [56] at turning points in the band structure clearly shown.

Charge density

The (pseudo) charge density for the n^{th} band is

$$\rho_n(\mathbf{r}) = e \sum_{\mathbf{k}} |\varphi_{n\mathbf{k}}(\mathbf{r})|^2, \tag{2.56}$$

where e is the electronic charge and the summation is over all states in a single Brillouin zone. By summing over the individual valence bands one obtains the valence charge density, illustrated in Figure 2.8. This shows both the covalent nature of the bonds, with the electrons most concentrated between the ions, and the polar nature of the InP crystal, with most of the valence charge around the higher valence group V (phosphorus) ion.

Dielectric function

The imaginary part of the dielectric function, per unit volume, is given by (see for example p31 of [25])

$$\epsilon_2(\omega) = \frac{\pi e^2}{m^2\omega^2\epsilon_0} \sum_{cv} \int |p_{cv}(\mathbf{k})|^2 \frac{d^3\mathbf{k}}{(2\pi)^3} \delta(\Delta E_{\mathbf{k}}), \tag{2.57}$$

where $\omega = E\hbar^{-1}$, $\Delta E_{\mathbf{k}} = (E_{\mathbf{k}} - E)$ and $p_{ij}(\mathbf{k})$ is the momentum matrix element between states in bands i and j at wave vector $\mathbf{k}$. The summation is such that only electrons excited from the valence to conduction bands ($v \rightarrow c$) are included, and the integration is over a single Brillouin zone.

[9]The sum is assumed to explicitly include the spin degeneracy of the bands. If this is not the case, a factor of two is required in the expression for $g(E)$.

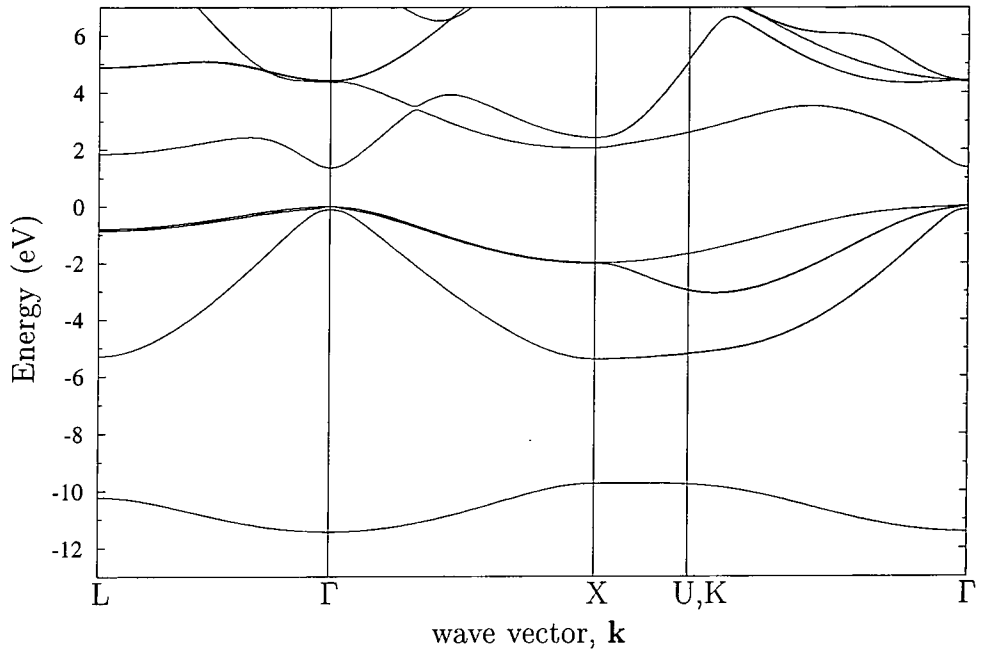


Figure 2.6 Electronic energy band structure of bulk indium phosphide, as calculated by the EPM. The material is modelled at room temperature using 89 plane waves per spin component.

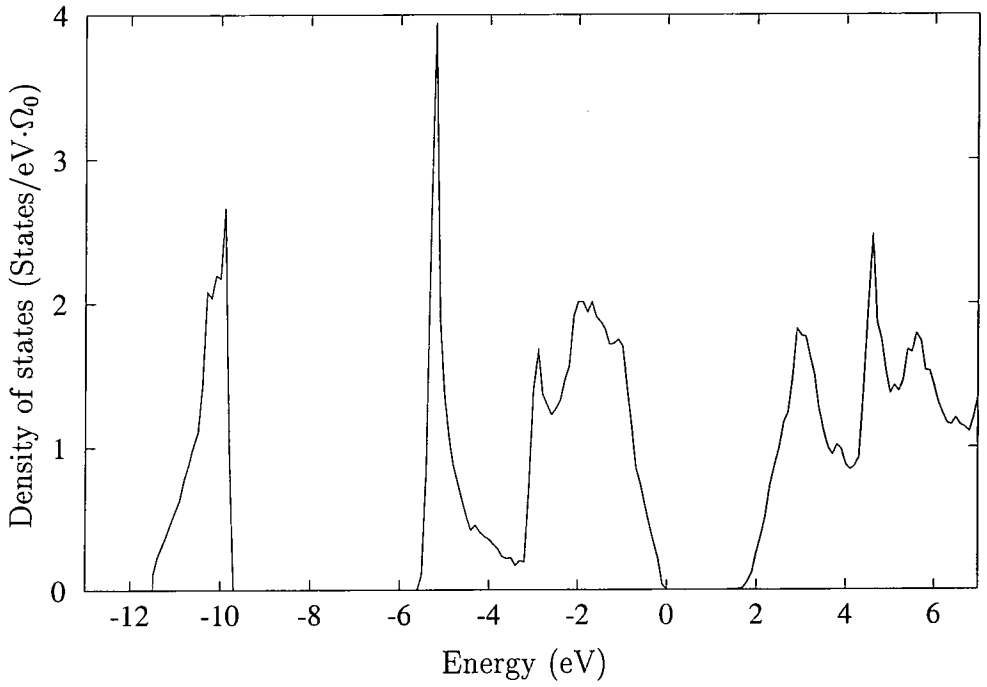


Figure 2.7 Electron density of states of indium phosphide, as calculated by the EPM. The material is modelled at room temperature using 89 plane waves per spin component. Ω_0 is the primitive unit cell volume.

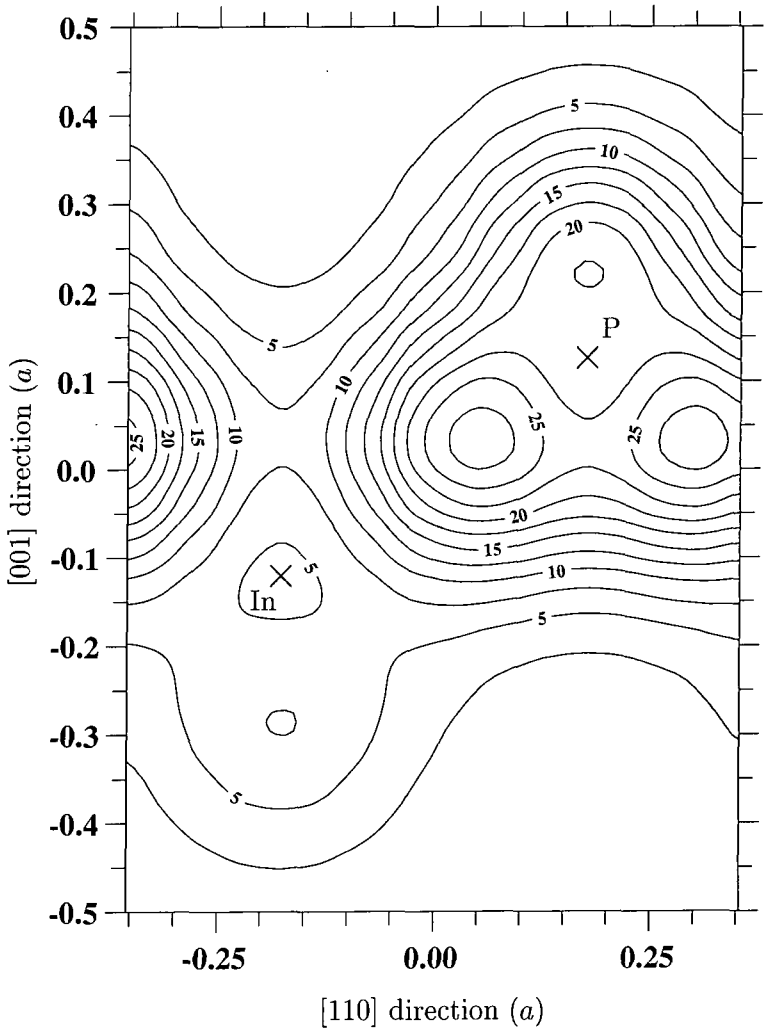


Figure 2.8 Total valence charge density for InP. The charge density contours are in units of electrons per unit cell volume. The (110) plane coincides with the covalent bond betwwen the two atoms shown, i.e., the origin of the plot lies at the bond midpoint. The position of the indium and phosphorus crystal ions are each indicated by $\times$.

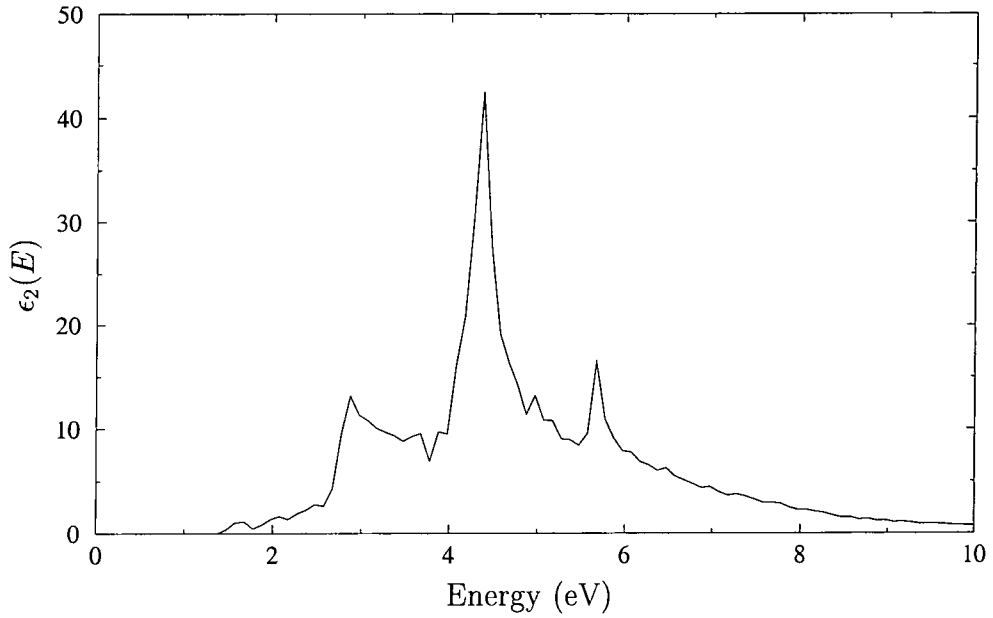


Figure 2.9 Imaginary part of the dielectric function for indium phosphide. The material is modelled at room temperature using 89 plane waves per spin component.

Figure 2.9 shows ϵ_2 for indium phosphide. Using the Kramers-Kronig relation it is possible to obtain the real part of the dielectric function from the imaginary part, and hence the reflectivity. This allows the comparison of calculated data with a directly measurable quantity. The relevant equations are given in Appendix B.

Ideally, one would use the reflectivity in the fitting process, i.e., adjust the form factors so that the calculated and experimental data is in agreement, since the experimentally determined transition energies used in the fitting are deduced from such data anyway. However, full analysis and comparison with experimental data of either the reflectivity (or, equivalently, $\epsilon_2(\omega)$) is extremely time consuming, since there are often many different transitions of varying strengths occurring in the same energy ranges. More importantly for the present work, the local pseudopotential method is known to be not as accurate in general as non-local methods [25] and, worse still, may introduce spurious structure in the data [51]. For these reasons the simpler approach of fitting to specific transitions is employed. The calculation of the dielectric function is used here to confirm that the bulk material parameters

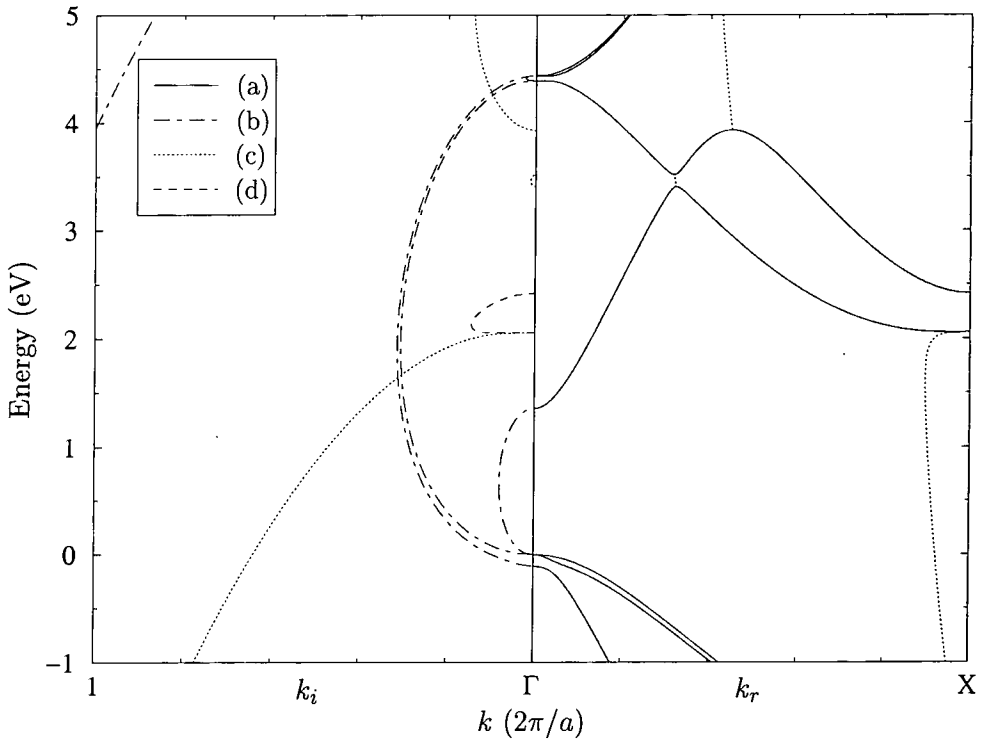


Figure 2.10 Complex band structure for indium phosphide at $\mathbf{k}_{||} = (0,0)$. The complex component, k, has real and imaginary parts k_r and k_i, respectively. The energy bands are labelled as follows: (a) purely real k, (b) purely imaginary k, (c) complex k, (d) complex k with real part at X. The material is modelled at room temperature using 89 plane waves per spin component.

produced by the fitting procedures generate reasonable energy band structure *and* wavefunctions throughout the zone, i.e., that $\epsilon_2(\omega)$ has the expected form and is in reasonable agreement with experimentally derived results.

Complex band structure

Although complex band structure is not relevant to the description of bulk materials (see Section 2.5), it is information that can be employed when a material forms a layer in a heterostructure, and so it makes sense to illustrate the nature of the complex band structure here. Figure 2.10 shows a plot of the complex band structure for indium phosphide. The data is for $\mathbf{k}_{||} = (0,0)$ and there is no strain present.

The details of complex band structure have been discussed by other workers

[54,57–61], here we point out the most important features for the present purposes. The most obvious feature of the complex band structure is the existence of states in the fundamental energy band gap. In particular, there is a loop in the imaginary plane joining the light hole and first conduction band, whilst the heavy hole and spin split off bands link up with the next conduction bands. In fact, the complex band structure links all turning points that have resulted from band anti-crossing, such as between the first and second conduction bands (at ≈ 0.3X and X for InP), in which case the loop has complex wave vector values, rather than the purely imaginary ones that loop over the band gap. It should also be pointed out that the energy bands may neither terminate nor coalesce [54], hence the number of complex-k solutions with real energy is constant, whatever the energy. When using a finite Fourier expansion of the wavefunction this can be used to check that the complex-k solutions have been correctly identified.

Chapter 3

Heterostructures — theory

3.1 Introduction

Chapter 2 outlined the particular form of the empirical pseudopotential method (EPM) that is used throughout this thesis. We now turn our attention to the application of the EPM to the analysis of heterostructures.

A heterostructure consists of one or more interfaces between two or more materials. The simplest heterostructure consists of a single interface between two different materials, the interface being called a heterojunction. By introducing further heterojunctions it is possible to create more complicated structures, such as quantum wells and potential barriers. In this work, all the interfaces that form heterojunctions are parallel (i.e., quantum wires and dots are not considered).

When a heterojunction is formed there is in general an offset between the valence and conduction band edges of each material (see [62,63] for reviews of this topic). This is illustrated in Figure 3.1, which shows the three types of band offset, or line up, considered in this work[1]. We assume that the interface is atomically abrupt and neglect band bending effects. This is done so that 'bulk' complex band structure

[1]Type III offsets, which are those between a semiconductor and semimetal, are not considered.

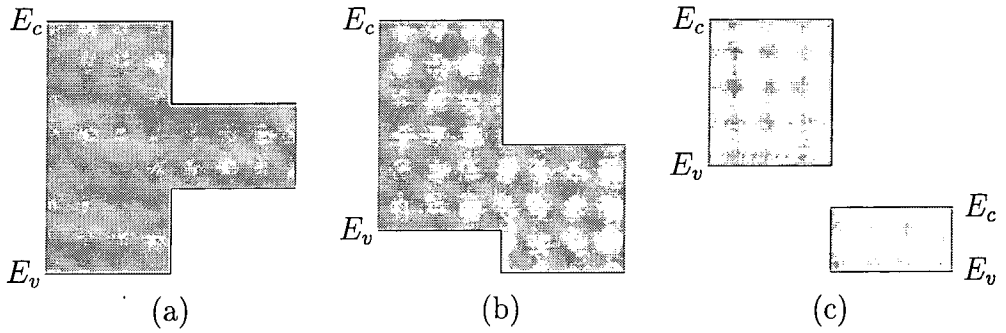


Figure 3.1 Schematic of the three types of heterojunction considered in this work. E_c and E_v label the conduction and valence band edges, respectively. (a) Type I (b) Type II — staggered (c) Type II — broken gap.

can be used in each material. We also assume that the materials are strained, when necessary, to form perfect, dislocation free interfaces, as discussed in Section 2.3.1. The band line up of a heterostructure consisting of two interfaces is shown in Figure 3.2. In this case, only two different materials are used, with a type I heterojunction

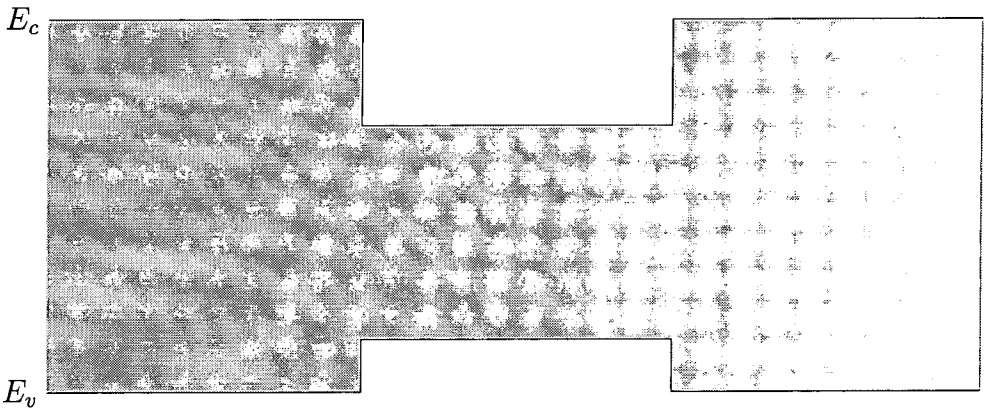


Figure 3.2 Schematic of a type I single quantum well (SQW). E_c and E_v label the conduction and valence band edges,respectively.

forming at the interface between them. Since the band gap of the central material is smaller than that of the material on either side, potential wells are formed in the conduction and valence bands. These are called quantum wells because the central layer width can be of the order of the electron de Broglie wave length and hence the system must be considered quantum mechanically.[2]

An electron in the conduction band well (or hole in the valence band well) with

[2]The layer width is also less than the mean free path of the electron.

energy below the conduction band edge of the barrier material is contained within the well. However, in the plane of the well, i.e., parallel to the interfaces, the periodicity of the bulk material is retained, hence the electron (hole) is free to move in this plane. Thus, the electron behaves in a quasi-two dimensional manner. Due to the electron's wave-like properties, only certain energies satisfy the quantum mechanical boundary conditions, hence the allowed energies are discrete, or quantised.

We are particularly interested in bound states in non-periodic heterostructures, i.e., those whose extreme left and right materials extend (theoretically) to $\mp\infty$. The definition of a bound state requires that the electron probability density tends to zero at infinity (see, for example, p29 of [64]), and we also have the usual condition that the probability density must remain finite at all spatial positions. These issues are discussed in more detail where applicable.

3.2 Interface boundary conditions

When calculating electron wavefunctions in heterostructures, it is vital to appreciate the correct boundary conditions that must be obeyed by the wavefunctions at the material interfaces (heterojunctions). The method used in this work for finding heterostructure states involves the joining at each interface of 'bulk' complex band structure states, calculated separately for each layer, in which case the boundary conditions are imposed to determine initially unknown expansion coefficients (see Section 3.3). It is important to realise that the boundary conditions are derived from the form of the (in this case, pseudo) Hamiltonian.

First, recall the form of the (pseudo) Schrödinger equation from (2.5), using the local EPM pseudopotential, (2.18),

$$\widehat{H^p}\varphi(\mathbf{r}) = \left(-\frac{\hbar^2}{2m_e}\nabla^2 + V^p(\mathbf{r})\right)\varphi(\mathbf{r}) = E\varphi(\mathbf{r}), \tag{3.1}$$

or,

$$-\frac{\hbar^2}{2m_e}\nabla^2\varphi(\mathbf{r}) = \big(E - V^p(\mathbf{r})\big)\varphi(\mathbf{r}). \tag{3.2}$$

Now, the right hand side of (3.2) is always finite, although in the present method used to model heterojunctions there is a discontinuity in the potential term at the interface due to the abrupt change in materials. Therefore, the left hand side must also be finite. However, the left hand side is the second spatial derivative of the wavefunction and for it to remain finite, the first derivative must be continuous. Requiring the first derivative to be continuous imposes continuity (and smoothness) on the wavefunction itself. To summarise,

- The wavefunction must be continuous over all space.
- The wavefunction gradient must be continuous over all space, i.e., the wavefunction must also be smooth over all space.

3.3 The transfer matrix method

The most general form of the electron wavefunction at energy E and in-plane wave vector $\mathbf{k}_{||}$ is a sum over all 'bulk' complex band structure states at that E and $\mathbf{k}_{||}$ [54,60], i.e., a sum over all the corresponding perpendicular wave vector components, k,

$$\Psi_{E,\mathbf{k}_{||}}(\mathbf{r}) = \sum_k C_{k,E,\mathbf{k}_{||}}\psi_{k,E,\mathbf{k}_{||}}(\mathbf{r}), \tag{3.3}$$

where the $\psi_{k,E,\mathbf{k}_{||}}(\mathbf{r})$ are wavefunctions, as in (2.51), and $C_{k,E,\mathbf{k}_{||}}$ are expansion coefficients to be determined. Certain values for the expansion coefficients allow the interface boundary conditions (see Section 3.2) to be satisfied, i.e., the boundary

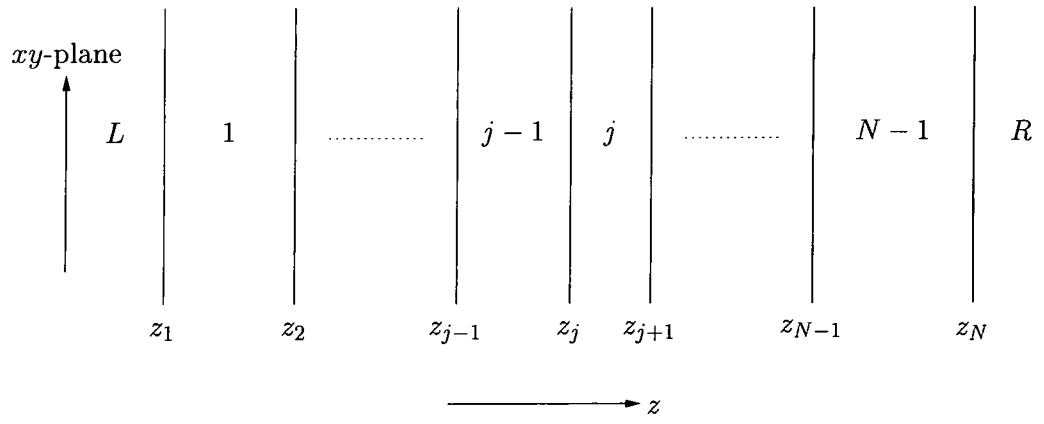


Figure 3.3 Schematic indicating N interfaces in a general heterostructure. The material layers are lablled L, 1, ... , $j-1$, j, ... , $N-1$ and R, where L and R are the left and right barriers, respectively.

conditions determine the expansion coefficients. For notational compactness, the implicit dependences on the energy and in-plane wave vector will not be indicated, unless their omission might cause confusion.

Now, consider a general heterojunction, labelled j, at position z_j, which joins layer $j-1$ to its left with layer j to its right, as in Figure 3.3. The boundary conditions require that

$$\Psi^{(j-1)}(z_j, \mathbf{r}_{\|}) = \Psi^{(j)}(z_j, \mathbf{r}_{\|}) \tag{3.4a}$$

$$\frac{\partial}{\partial z}\Psi^{(j-1)}\bigg|_{z_j,\mathbf{r}_{\|}} = \frac{\partial}{\partial z}\Psi^{(j)}\bigg|_{z_j,\mathbf{r}_{\|}}, \tag{3.4b}$$

where the superscripts in parenthesis are used to denote the material layers. So, first matching the wavefunctions, using (2.27), we have

$$\sum_k C_k^{(j-1)} \sum_{\mathbf{g}} \begin{pmatrix} a_{\mathbf{g}k\uparrow}^{(j-1)} \\ a_{\mathbf{g}k\downarrow}^{(j-1)} \end{pmatrix} e^{i(k+g)z_j} e^{i(\mathbf{k}_{\|}+\mathbf{g}_{\|})\cdot\mathbf{r}_{\|}}$$

$$= \sum_{k'} C_{k'}^{(j)} \sum_{\mathbf{g}'} \begin{pmatrix} a_{\mathbf{g}'k'\uparrow}^{(j)} \\ a_{\mathbf{g}'k'\downarrow}^{(j)} \end{pmatrix} e^{i(k'+g')z_j} e^{i(\mathbf{k}'_{\|}+\mathbf{g}'_{\|})\cdot\mathbf{r}_{\|}}, \tag{3.5}$$

with the complex k from the appropriate layers being used on each side of the equation, i.e., $k = k^{(j-1)}$ and $k = k^{(j)}$ on the left and right hand sides, respectively. Since there is a common periodicity in the plane of the interface, momentum conservation requires $\mathbf{k}_{||} = \mathbf{k}'_{||}$. Next, we multiply by $e^{-i\mathbf{g}''_{||}\cdot\mathbf{r}_{||}}$ and integrate over a unit cell face in the plane of the interface to obtain

$$\sum_k C_k^{(j-1)} \sum_{\mathbf{g}} \begin{pmatrix} a^{(j-1)}_{\mathbf{g}k\uparrow} \\ a^{(j-1)}_{\mathbf{g}k\downarrow} \end{pmatrix} e^{i(k+g)z_j} \delta_{\mathbf{g}_{||}\mathbf{g}''_{||}} = \sum_{k'} C_{k'}^{(j)} \sum_{\mathbf{g}'} \begin{pmatrix} a^{(j)}_{\mathbf{g}'k'\uparrow} \\ a^{(j)}_{\mathbf{g}'k'\downarrow} \end{pmatrix} e^{i(k'+g')z_j} \delta_{\mathbf{g}'_{||}\mathbf{g}''_{||}}, \quad (3.6)$$

which equates the coefficients of each two dimensional (2D) projection of the wavefunctions onto the interface [21].

This can be further simplified if z_j is an integer number of monolayers[3] [45]. The zinc-blende crystal has an underlying face centred cubic Bravais lattice and hence a body centred cubic reciprocal lattice and therefore the components of the reciprocal lattice vectors, $\mathbf{g}$, are either all even or all odd (in units of $2\pi/a$). It follows that the Kronecker delta functions in (3.6) ensure that both g and g' will have the same parity as the components of $\mathbf{g}''_{||}$, i.e., $g = m2\pi/a$ and $g' = m'2\pi/a$, where $\mod(m,2) = \mod(m',2)$ and $m, m' \in \mathbb{Z}$. So, restricting z_j to an integer number of monolayers, $z_j = na/2$ with $n \in \mathbb{Z}$, gives

$$e^{igz_j} = e^{i\pi mn} = \begin{cases} -1 & mn \text{ odd} \\ +1 & mn \text{ even}, \end{cases} \quad (3.7)$$

[3] A monolayer is defined as half a lattice parameter, i.e., *two* atomic layers. This constraint on the layer widths is relaxed slightly in Chapters 5 and 6

and

$$e^{ig'z_j} = e^{i\pi m'n} = \begin{cases} -1 & m'n \text{ odd} \\ +1 & m'n \text{ even.} \end{cases} \tag{3.8}$$

Therefore, since m and m' are both odd or both even, e^{igz_j} and $e^{ig'z_j}$ will always cancel, and (3.6) becomes [21]

$$\sum_k C_k^{(j-1)} P_{\mathbf{g}_\parallel k}^{(j-1,r)} = \sum_{k'} C_{k'}^{(j)} P_{\mathbf{g}_\parallel k'}^{(j,l)}, \tag{3.9}$$

where

$$P_{\mathbf{g}_\parallel k}^{(j,r)} = \sum_{\mathbf{g}'} \begin{pmatrix} a_{\mathbf{g}'k\uparrow}^{(j)} \\ a_{\mathbf{g}'k\downarrow}^{(j)} \end{pmatrix} e^{ikz_{j+1}} \delta_{\mathbf{g}'_\parallel \mathbf{g}_\parallel} \tag{3.10a}$$

$$P_{\mathbf{g}_\parallel k}^{(j,l)} = \sum_{\mathbf{g}'} \begin{pmatrix} a_{\mathbf{g}'k\uparrow}^{(j)} \\ a_{\mathbf{g}'k\downarrow}^{(j)} \end{pmatrix} e^{ikz_j} \delta_{\mathbf{g}'_\parallel \mathbf{g}_\parallel}. \tag{3.10b}$$

$P_{\mathbf{g}_\parallel k}^{(j,r)}$ gives the $\mathbf{g}_\parallel$ projection of the wavefunction with complex wave vector k in layer j onto the interface to its right, i.e., the $(j+1)^{th}$ interface at z_{j+1}. $P_{\mathbf{g}_\parallel k}^{(j,l)}$ gives the $\mathbf{g}_\parallel$ projection of the wavefunction with complex wave vector k in layer j onto the interface to its left, i.e., the j^{th} interface at z_j.

The same procedure is followed for matching the gradient of the wavefunction at the interface, leading to

$$\sum_k C_k^{(j-1)} Q_{\mathbf{g}_\parallel k}^{(j-1,r)} = \sum_{k'} C_{k'}^{(j)} Q_{\mathbf{g}_\parallel k'}^{(j,l)}, \tag{3.11}$$

where

$$Q^{(j,r)}_{\mathbf{g}_{\|}k} = \sum_{\mathbf{g}'}(k+g') \begin{pmatrix} a^{(j)}_{\mathbf{g}'k\uparrow} \\ a^{(j)}_{\mathbf{g}'k\downarrow} \end{pmatrix} e^{ikz_{j+1}} \delta_{\mathbf{g}'_{\|}\mathbf{g}_{\|}} \tag{3.12a}$$

$$Q^{(j,l)}_{\mathbf{g}_{\|}k} = \sum_{\mathbf{g}'}(k+g') \begin{pmatrix} a^{(j)}_{\mathbf{g}'k\uparrow} \\ a^{(j)}_{\mathbf{g}'k\downarrow} \end{pmatrix} e^{ikz_{j}} \delta_{\mathbf{g}'_{\|}\mathbf{g}_{\|}}. \tag{3.12b}$$

We now combine (3.9) and (3.11) in the form

$$\begin{pmatrix} \mathbf{P}^{(j-1,r)} \\ \mathbf{Q}^{(j-1,r)} \end{pmatrix} \mathbf{C}^{(j-1)} = \begin{pmatrix} \mathbf{P}^{(j,l)} \\ \mathbf{Q}^{(j,l)} \end{pmatrix} \mathbf{C}^{(j)}, \tag{3.13}$$

where

$$\mathbf{P}^{(j,l)} = \begin{pmatrix} P^{(j,l)}_{(00)\uparrow k_1} & P^{(j,l)}_{(00)\uparrow k_2} & \cdots \\ P^{(j,l)}_{(11)\uparrow k_1} & P^{(j,l)}_{(11)\uparrow k_2} & \cdots \\ \vdots & \vdots & \\ P^{(j,l)}_{(00)\downarrow k_1} & P^{(j,l)}_{(00)\downarrow k_2} & \cdots \\ P^{(j,l)}_{(11)\downarrow k_1} & P^{(j,l)}_{(11)\downarrow k_2} & \cdots \\ \vdots & \vdots & \end{pmatrix} \qquad \mathbf{C}^{(j)} = \begin{pmatrix} C^{(j)}_{k_1} \\ C^{(j)}_{k_2} \\ \vdots \end{pmatrix} \tag{3.14}$$

with $\mathbf{P}^{(j-1,r)}$, $\mathbf{Q}^{(j-1,r)}$, $\mathbf{Q}^{(j,l)}$ and $\mathbf{C}^{(j-1)}$ being defined in analogous fashion. Now, it was pointed out in Section 2.5 that the number of non-equivalent complex-k solutions is four times the number of 2D reciprocal lattice vectors, $\mathbf{g}_{\|}$, if spin is included. Suppose there are M $\mathbf{g}_{\|}$'s, then the matrices of the type $\mathbf{P}^{(j,l)}$ will be of dimension $2M \times 4M$ with $\mathbf{C}^{(j)}$ being $4M \times 1$, and so those in (3.13) will be $4M \times 4M$,

thus we may write

$$\mathbf{C}^{(j)} = \begin{pmatrix} \mathbf{P}^{(j,l)} \\ \mathbf{Q}^{(j,l)} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{P}^{(j-1,r)} \\ \mathbf{Q}^{(j-1,r)} \end{pmatrix} \mathbf{C}^{(j-1)} \tag{3.15}$$

or,

$$\mathbf{C}^{(j)} = \mathbf{T}^{(j)}\mathbf{C}^{(j-1)} \tag{3.16}$$

which defines the widely used *transfer matrix* (see, for example, [48]), $\mathbf{T}^{(j)}$, relating the wavefunctions on either side of the j^{th} interface. The advantage of defining a transfer matrix becomes clear when there are many layers in the heterostructure. Since (3.16) is quite general (apart from the exceptions explained below), we may also write

$$\mathbf{C}^{(j+1)} = \mathbf{T}^{(j+1)}\mathbf{C}^{(j)} = \mathbf{T}^{(j+1)}\mathbf{T}^{(j)}\mathbf{C}^{(j-1)}, \text{ etc,} \tag{3.17}$$

and so the wavefunctions from any two layers, j and $j+N$, may be related by

$$\mathbf{C}^{(j+N)} = \mathbf{T}^{(j+N)}\mathbf{T}^{(j+N-1)} \ldots \mathbf{T}^{(j+2)}\mathbf{T}^{(j+1)}\mathbf{C}^{(j)}. \tag{3.18}$$

It should be noted that this analysis is only valid for interfaces between material layers of finite width, since it requires that all the complex-k solutions are included in the matching[4]. The two semi-infinite barrier layers are treated separately (see Section 3.4).

[4]Strictly, we require 4 times the number of $\mathbf{g}_{\|}$'s that are used in the matching.

3.4 Bound states

We now apply the transfer matrix method to calculating heterostructure bound states, i.e., those states that are essentially localised in some finite range of coordinate z. For this to be the case, we require that the integral of the probability density over all space is finite,

$$\int_{\text{all } \mathbf{r}} |\Psi(\mathbf{r})|^2 \, d^3\mathbf{r} \Rightarrow \text{finite}. \tag{3.19}$$

Hence, (3.19) must be modified accordingly to read

$$\int_{z=-\infty}^{z=+\infty} \int_{\Omega_{\|}} |\Psi(\mathbf{r})|^2 \, d^3\mathbf{r} = \text{finite}, \tag{3.20}$$

where it is assumed that the interfaces are in the xy-plane and the xy-component of the wavefunction is normalised over some suitable in-plane area $\Omega_{\|}$, normally the conventional unit cell face.

The general expression for the wavefunction in any layer is given by (3.3), but from (3.20) it is clear that bound states must be decaying in the left and right barriers and therefore *all* the 'bulk' states, ψ, in the sum (3.3) must have a non-zero imaginary wave vector component in the perpendicular direction. In the left (right) barrier the wavefunction must decay with decreasing (increasing) distance, so, since $\psi \propto e^{i\mathbf{k}\cdot\mathbf{r}}$, we have the following conditions for complex wave vector component, $k = k_r + ik_i$,

$$\text{Left barrier:} \quad \lim_{z\to-\infty} \left(e^{i(k_r+ik_i)z}\right) = \lim_{z\to-\infty} \left(e^{-k_i z} e^{ik_r z}\right) = 0, \quad \text{if} \quad k_i < 0, \tag{3.21a}$$

$$\text{Right barrier:} \quad \lim_{z\to+\infty} \left(e^{i(k_r+ik_i)z}\right) = \lim_{z\to+\infty} \left(e^{-k_i z} e^{ik_r z}\right) = 0, \quad \text{if} \quad k_i > 0. \tag{3.21b}$$

So, not only is the energy range in which bound states can exist limited by the

requirement that all the complex wave vectors in the left and right barriers have non-zero imaginary parts, but in these layers, exactly half the 'bulk' states are not allowed in the expression for the total wavefunction, (3.3).

The boundary conditions applied at the first and last interfaces, 1 and N, are as for a general interface, i.e., those in (3.4). So, at the first interface, the wavefunctions from the left barrier, L, and the first finite width layer, layer 1, are matched by

$$\sum_{k=k^{(-)}} C_k^{(L)} \sum_{\mathbf{g}} \begin{pmatrix} a_{\mathbf{g}k\uparrow}^{(L)} \\ a_{\mathbf{g}k\downarrow}^{(L)} \end{pmatrix} e^{i(k+g)z_1} e^{i(\mathbf{k}_\parallel+\mathbf{g}_\parallel)\cdot\mathbf{r}_\parallel} = \sum_{k'=k^{(1)}} C_{k'}^{(1)} \sum_{\mathbf{g}'} \begin{pmatrix} a_{\mathbf{g}'k'\uparrow}^{(1)} \\ a_{\mathbf{g}'k'\downarrow}^{(1)} \end{pmatrix} e^{i(k'+g')z_1} e^{i(\mathbf{k}'_\parallel+\mathbf{g}'_\parallel)\cdot\mathbf{r}_\parallel}, \quad (3.22)$$

where $k^{(-)}$ indicates a complex wave vector with a negative imaginary part. Similarly, at the final interface the wavefunctions are matched with

$$\sum_{k=k^{(+)}} C_k^{(R)} \sum_{\mathbf{g}} \begin{pmatrix} a_{\mathbf{g}k\uparrow}^{(R)} \\ a_{\mathbf{g}k\downarrow}^{(R)} \end{pmatrix} e^{i(k+g)z_N} e^{i(\mathbf{k}_\parallel+\mathbf{g}_\parallel)\cdot\mathbf{r}_\parallel} = \sum_{k'=k^{(N-1)}} C_{k'}^{(N-1)} \sum_{\mathbf{g}'} \begin{pmatrix} a_{\mathbf{g}'k'\uparrow}^{(N-1)} \\ a_{\mathbf{g}'k'\downarrow}^{(N-1)} \end{pmatrix} e^{i(k'+g')z_N} e^{i(\mathbf{k}'_\parallel+\mathbf{g}'_\parallel)\cdot\mathbf{r}_\parallel}, \quad (3.23)$$

where $k^{(+)}$ indicates a complex wave vector with a positive imaginary part. The expressions for the matched wavefunction gradients follow in a natural way, cf. (3.11) and (3.12), giving the following equations for each 2D reciprocal lattice vector

projection

$$\sum_{k=k^{(-)}} C_k^{(L)} P_{\mathbf{g}_\parallel k}^{(L)} = \sum_{k'=k^{(1)}} C_{k'}^{(1)} P_{\mathbf{g}_\parallel k'}^{(1,l)} \tag{3.24a}$$

$$\sum_{k=k^{(-)}} C_k^{(L)} Q_{\mathbf{g}_\parallel k}^{(L)} = \sum_{k'=k^{(1)}} C_{k'}^{(1)} Q_{\mathbf{g}_\parallel k'}^{(1,l)} \tag{3.24b}$$

$$\sum_{k=k^{(+)}} C_k^{(R)} P_{\mathbf{g}_\parallel k}^{(R)} = \sum_{k'=k^{(N-1)}} C_{k'}^{(N-1)} P_{\mathbf{g}_\parallel k'}^{(N-1,r)} \tag{3.24c}$$

$$\sum_{k=k^{(+)}} C_k^{(R)} P_{\mathbf{g}_\parallel k}^{(R)} = \sum_{k'=k^{(N-1)}} C_{k'}^{(N-1)} P_{\mathbf{g}_\parallel k'}^{(N-1,r)}. \tag{3.24d}$$

We now use the transfer matrix method to couple the wavefunction in the final finite width layer, $N-1$, to that in the first, i.e., from (3.16),

$$\mathbf{C}^{(N-1)} = \mathbf{T}^{(N-1)}\mathbf{T}^{(N-2)} \cdots \mathbf{T}^{(3)}\mathbf{T}^{(2)}\mathbf{C}^{(1)} = \mathbf{T}\mathbf{C}^{(1)}, \tag{3.25}$$

which defines the 'total' transfer matrix, $\mathbf{T}$, for the entire heterostructure. Now, exactly half of the $4M$ complex-k solutions are excluded from the left hand side of (3.24), so the four equations in (3.24) can be combined into two matrix equations, using (3.25),

$$\begin{pmatrix} \mathbf{P}^{(L)} & 0 \\ 0 & \mathbf{P}^{(R)} \end{pmatrix} \begin{pmatrix} \mathbf{C}^{(L)} \\ \mathbf{C}^{(R)} \end{pmatrix} = \begin{pmatrix} \mathbf{P}^{(1,l)} \\ \mathbf{P}^{(N-1,r)}\mathbf{T} \end{pmatrix} \mathbf{C}^{(1)} \tag{3.26}$$

$$\begin{pmatrix} \mathbf{Q}^{(L)} & 0 \\ 0 & \mathbf{Q}^{(R)} \end{pmatrix} \begin{pmatrix} \mathbf{C}^{(L)} \\ \mathbf{C}^{(R)} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}^{(1,l)} \\ \mathbf{Q}^{(N-1,r)}\mathbf{T} \end{pmatrix} \mathbf{C}^{(1)}, \tag{3.27}$$

or, in more compact notation

$$\mathbf{P}^{(LR)}\mathbf{C}^{(LR)} = \mathbf{P}^{(1:N-1)}\mathbf{C}^{(1)} \tag{3.28a}$$

$$\mathbf{Q}^{(LR)}\mathbf{C}^{(LR)} = \mathbf{Q}^{(1:N-1)}\mathbf{C}^{(1)}. \tag{3.28b}$$

Since $\mathbf{P}^{(LR)}$, $\mathbf{Q}^{(1:N-1)}$, $\mathbf{P}^{(1:N-1)}$ and $\mathbf{Q}^{(1:N-1)}$ are square matrices (of order $4M$), the equations (3.28) may be combined into a single eigenequation, with eigenvalue 1 and eigenvector $\mathbf{C}^{(1)}$ [21],

$$\left(\mathbf{Q}^{(1:N-1)}\right)^{-1}\mathbf{Q}^{(LR)}\left(\mathbf{P}^{(LR)}\right)^{-1}\mathbf{P}^{(1:N-1)}\mathbf{C}^{(1)} = \mathbf{M}^{(1)}\mathbf{C}^{(1)} = \mathbf{C}^{(1)}. \tag{3.29}$$

So, bound states of the heterostructure exist for those combinations of energy and in-plane wave vector resulting in complex-k at which (3.29) is satisfied, i.e., when one of the eigenvalues of $\mathbf{M}^{(1)}$ is unity. The corresponding eigenvector then gives the expansion coefficients for the total wavefunction in layer 1, from which all the remaining coefficients may be calculated, using (3.16) for the other finite width layers and (3.28) for the left and right barriers.

Equivalently, non-trivial solutions to (3.29) exist only when the determinant of $\mathbf{M}^{(1)} - \mathbf{I}$ vanishes, where $\mathbf{I}$ is the identity matrix, since

$$\left(\mathbf{M}^{(1)} - \mathbf{I}\right)\mathbf{C}^{(1)} = 0, \tag{3.30}$$

and therefore bound states exist when

$$\left|\mathbf{M}^{(1)} - \mathbf{I}\right| = 0. \tag{3.31}$$

So, for a given in-plane wave vector, we scan through the appropriate energy range calculating the complex band structure and Fourier coefficients for each layer at each energy step, testing to see if condition (3.29) and/or (3.31) holds, in which case the energy of a bound state has been found.

3.5 Momentum matrix elements

From Fermi's Golden Rule, the transition rate between two electronic states induced by electromagnetic (EM) radiation is dependent on the size of the squared magnitude of the momentum matrix elements between the states (see, for example, section 13.6 of [1]). The relevant equations for calculating momentum matrix elements in a heterostructure using the EPM is outlined below.

The momentum matrix element in the direction parallel to the unit vector $\mathbf{e}$, between two states i and j, is given by

$$p_{ij} = \left\langle \Psi^{(i)} \middle| \mathbf{e} \cdot \widehat{\mathbf{p}} \middle| \Psi^{(j)} \right\rangle, \tag{3.32}$$

where $\widehat{\mathbf{p}} = -i\hbar\nabla$ is the quantum mechanical momentum operator. For bound states in heterostructures, (3.32) becomes

$$p_{ij} = \int_{z=-\infty}^{z=+\infty} \int_{\Omega_{||}} \left(\Psi^{(i)}\right)^* \mathbf{e} \cdot \widehat{\mathbf{p}} \Psi^{(j)} d^3\mathbf{r}, \tag{3.33}$$

giving the momentum matrix element per unit cell area. In practice, the contribution to p_{ij} from each layer of the heterostructure is calculated separately, then summed to get the final value. So, for a heterostructure made up of $N-1$ layers of finite width, we use

$$p_{ij} = p_{ij}^{(L)} + p_{ij}^{(1)} + \cdots + p_{ij}^{(N-1)} + p_{ij}^{(R)}, \tag{3.34}$$

with the contribution from the l^{th} layer given by

$$p_{ij}^{(l)} = \int_{z_{\min}^{(l)}}^{z_{\max}^{(l)}} \int_{\Omega_{||}} \left[\sum_{k\mathbf{g}} \left(C_k^{(l)}\right)^* \begin{pmatrix} a_{k\mathbf{g}\uparrow}^{(l)} \\ a_{k\mathbf{g}\downarrow}^{(l)} \end{pmatrix}^{\dagger} e^{-i(\mathbf{k}^*+\mathbf{g})\cdot\mathbf{r}} \right.$$

$$\left. \times \mathbf{e}\cdot\widehat{\mathbf{p}} \sum_{k'\mathbf{g}'} C_{k'}^{(l)} \begin{pmatrix} a_{k'\mathbf{g}'\uparrow}^{(l)} \\ a_{k'\mathbf{g}'\downarrow}^{(l)} \end{pmatrix} e^{-i(\mathbf{k}'+\mathbf{g}')\cdot\mathbf{r}} \right] d^3\mathbf{r} \tag{3.35}$$

$$= \hbar\Omega_{||} \sum_{kk'\mathbf{g}\mathbf{g}'} \sum_{ss'} \left(C_k^{(l)}\right)^* C_{k'}^{(l)} \left(a_{k\mathbf{g}s}^{(l)}\right)^* a_{k'\mathbf{g}'s'}\, \mathbf{e}\cdot(\mathbf{k}'+\mathbf{g}')\, \delta_{\mathbf{g}\mathbf{g}'}\delta_{ss'}$$

$$\times I^{(l)}, \tag{3.36}$$

where

$$I^{(l)} = \int_{z_{\min}^{(l)}}^{z_{\max}^{(l)}} e^{i(k'-k^*+g'-g)z} dz. \tag{3.37}$$

This integral can be calculated with the following results, depending on the layer type,

$$I^{(L)} = e^{i\alpha z_L}/(i\alpha) \tag{3.38}$$

$$I^{(w)} = \begin{cases} (e^{i\alpha z_{\max}^{(w)}} - e^{i\alpha z_{\min}^{(w)}})/(i\alpha) & \alpha \neq 0 \\ z_{\max}^{(w)} - z_{\min}^{(w)} & \alpha = 0 \end{cases} \tag{3.39}$$

$$I^{(R)} = -e^{i\alpha z_R}/(i\alpha) \tag{3.40}$$

where $\alpha = (k' - k^* + g' - g)$ and w indicates a general layer of finite width between the extreme left and right barriers (L and R), with barrier interfaces at z_L and z_R (labelled z_1 and z_N in Figure 3.3).

3.6 Dipole matrix elements

A basic principle of electromagnetic (EM) theory is that the vector and scalar potentials are not uniquely defined for given electric and magnetic fields. Thus it is possible to alter the potentials in certain ways, so-called *gauge transformations*, without affecting the fields (see, for example, p132 of [65] or p536 of [66]). This is called *gauge invariance*. When including the interaction between a charged particle and an electromagnetic field, as in Fermi's Golden Rule, gauge transformations do affect the form of the Hamiltonian and, hence, the wavefunction of the particle, whilst gauge invariance ensures that all observables are unchanged.

In Section 3.5 we stated that electron transitions induced by electromagnetic radiation take place at a rate dependent on the squared magnitude of the momentum matrix element between the states. This is correct, but we implicitly assumed that the term in the Hamiltonian describing the electron-field interaction was of the form $\mathbf{A} \cdot \widehat{\mathbf{p}}$, where $\mathbf{A}$ is the vector potential of the electromagnetic field. However, due to gauge invariance, this is just one form of the Hamiltonian, specifically that assuming the *Coulomb gauge*, which defines $\nabla \cdot \mathbf{A} = 0$. By making the appropriate gauge transformation, we arrive at an alternative Hamiltonian, where the term that includes the interaction between the particle and the EM field is of the form $\mathbf{E} \cdot \mathbf{r}$, where $\mathbf{E}$ is the electric field component of the EM radiation (see, for example, p551 of [66]). In this gauge, transition rates are characterised by the so-called dipole matrix elements between the states involved. Thus, to obtain transition rate information, one may calculate either momentum or dipole matrix elements, according to preference and convenience.

The dipole matrix element between states i and j, along the direction of the unit

vector $\mathbf{e}$, is given by

$$r_{ij} = e\langle \Psi^{(i)} | \mathbf{e} \cdot \mathbf{r} | \Psi^{(j)} \rangle, \tag{3.41}$$

where e is the charge on the electron. For bound states in heterostructures, (3.41) becomes

$$r_{ij} = e \int_{z=-\infty}^{z=+\infty} \int_{\Omega_{\|}} \left(\Psi^{(i)}\right)^* \mathbf{e} \cdot \mathbf{r} \Psi^{(j)} d^3\mathbf{r}, \tag{3.42}$$

giving the dipole matrix element per unit cell area. In this thesis, we are interested in dipole matrix elements, z_{ij}, along the growth direction, z. The contribution to z_{ij} from each layer of the heterostructure is calculated separately, then summed to get the final value. So, for a heterostructure made up of $N-1$ layers of finite width, we use

$$z_{ij} = z_{ij}^{(L)} + z_{ij}^{(1)} + \cdots + z_{ij}^{(N-1)} + z_{ij}^{(R)}. \tag{3.43}$$

The contribution from the l^{th} layer is given by

$$\begin{aligned} z_{ij}^{(l)} = e \int_{z_{\min}^{(l)}}^{z_{\max}^{(l)}} \int_{\Omega_{\|}} \Bigg[\sum_{\mathbf{kg}} \left(C_k^{(l)}\right)^* \begin{pmatrix} a_{\mathbf{kg}\uparrow}^{(l)} \\ a_{\mathbf{kg}\downarrow}^{(l)} \end{pmatrix}^{\dagger} e^{-i(\mathbf{k}^*+\mathbf{g})\cdot\mathbf{r}} \\ \times z \sum_{\mathbf{k'g'}} C_{k'}^{(l)} \begin{pmatrix} a_{\mathbf{k'g'}\uparrow}^{(l)} \\ a_{\mathbf{k'g'}\downarrow}^{(l)} \end{pmatrix} e^{-i(\mathbf{k'}+\mathbf{g'})\cdot\mathbf{r}} \Bigg] d^3\mathbf{r} \end{aligned} \tag{3.44}$$

$$\begin{aligned} = e\Omega_{\|} \sum_{\mathbf{kk'gg'}} \sum_{ss'} \left(C_k^{(l)}\right)^* C_{k'}^{(l)} \left(a_{\mathbf{kg}s}^{(l)}\right)^* a_{\mathbf{k'g'}s'} \delta_{\mathbf{gg'}} \delta_{ss'} \\ \times I_z^{(l)}, \end{aligned} \tag{3.45}$$

where

$$I_z^{(l)} = \int_{z_{\min}^{(l)}}^{z_{\max}^{(l)}} z e^{i(k'-k^*+g-g')z} dz. \tag{3.46}$$

This integral can be solved on integration by parts to yield

$$I_z^{(L)} = (1 - i\alpha z_L) e^{i\alpha z_L} \alpha^{-2} \tag{3.47}$$

$$I_z^{(w)} = \begin{cases} \left[(1 - i\alpha z_{\max}^{(w)}) e^{i\alpha z_{\max}^{(w)}} - (1 - i\alpha z_{\min}^{(w)}) e^{i\alpha z_{\min}^{(w)}}\right] \alpha^{-2} & \alpha \neq 0 \\ ({z_{\max}^{(w)}}^2 - {z_{\min}^{(w)}}^2)/2 & \alpha = 0 \end{cases} \tag{3.48}$$

$$I_z^{(R)} = -(1 - i\alpha z_R) e^{i\alpha z_R} \alpha^{-2} \tag{3.49}$$

where the symbols have the same meanings as in the momentum matrix elements calculation, Section 3.5.

3.7 Joint density of states

The transition rate caused by EM radiation at a certain frequency or corresponding photon energy depends not just on the matrix elements between the states involved, but also on the number of both initial and final states, i.e., the joint density of states (JDOS) in energy. We label the JDOS at energy E, between bands i and j, by $g_{ij}(E)$. The full JDOS at energy E is then given by summing the contributions from all the $g_{ij}(E)$, i.e., from all transitions[5] of energy E.

Since the heterostructures under consideration are quasi-2D systems, the underlying form of $g_{ij}(E)$ is step-like (see, for example, section 8.8 of [1]). The ideal textbook form assumes parabolic in-plane energy bands, $E_{\mathbf{k}_\parallel} \propto \mathbf{k}_\parallel^2$, which is an increasingly inaccurate approximation as the in-plane wave vector, $\mathbf{k}_\parallel$, moves away

[5]Requiring the initial band to be occupied prevents the sum from including a theoretically infinite number of band-to-band transitions in the conduction band.

from the 2D Brillouin zone ($BZ_{||}$) centre, $\mathbf{k}_{||} = (0,0)$. Hence, an explicit calculation of $g_{ij}(E)$ is necessary.[6]

The two dimensional joint density of states may be expressed as

$$g_{ij}(E) = \frac{2}{(2\pi)^2} \int_{\ell(E)} \frac{d\ell}{|\nabla_{\mathbf{k}_{||}} E_{\mathbf{k}_{||}}|}, \tag{3.50}$$

where $\ell(E)$ is the line in $\mathbf{k}_{||}$-space within $BZ_{||}$ with energy E, $d\ell$ is an infinitessimal distance along this line, and $E_{\mathbf{k}_{||}} = E^{(j)}_{\mathbf{k}_{||}} - E^{(i)}_{\mathbf{k}_{||}}$ is the energy separation of the two bands. An exact evaluation of this expression requires full knowledge of the in-plane band structure of the two states so that $E_{\mathbf{k}_{||}}$ and then $|\nabla_{\mathbf{k}_{||}} E_{\mathbf{k}_{||}}|$ may be obtained. Calculation of the in-plane band structure is extremely time consuming when using the empirical pseudopotential method. Therefore, an exact evaluation of (3.50) is not feasible. Fortunately, if the in-plane band structure is close to isotropic in $\mathbf{k}_{||}$-space[7] we may make the approximation $E_{\mathbf{k}_{||}} \approx E_{|\mathbf{k}_{||}|} = E_{k_{||}}$ to obtain

$$\left|\nabla_{\mathbf{k}_{||}} E_{\mathbf{k}_{||}}\right| \rightarrow \left|\nabla_{k_{||}} E_{k_{||}}\right|. \tag{3.51}$$

This means that the path $\ell(E)$ is approximated by a circle, so that

$$\int_{\ell(E)} d\ell \approx 2\pi k_{||}(E), \tag{3.52}$$

with $k_{||}(E)$ being the value of $k_{||}$ at energy E (if there is more than one $k_{||}$ with energy E, i.e., there is more than a single $\ell(E)$, we must sum over each case). Substituting (3.51) and (3.52) into (3.50) leads to the expression actually used to calculate the

[6] In any case, some subbands may not be even approximately parabolic.

[7] This is the case for the InAs/AlSb conduction band states considered in Chapter 4.

JDOS in this work [67],

$$g_{ij}(E) = \frac{k_{\|}(E)}{\pi |\nabla_{k_{\|}} E_{k_{\|}}|}. \tag{3.53}$$

This approximation has the advantage that we need only calculate the in-plane band structure along a single line in $\mathbf{k}_{\|}$-space in order to obtain the data necessary for determining the JDOS[8].

3.8 Applied electric field

Many device applications of heterostructures require an electric field normal to the well plane (parallel to the z-axis in this work). The application of an electric field modifies the energy band structure of the system, since the associated potential energy is a perturbation to the Hamiltonian in the absence of the field.

In the present work, we restrict the analysis to cases involving uniform fields, such as those resulting from applying a bias across the different layers of a heterostructure. From basic electrostatic theory (see, for example, p4 of [68]), the electric field, $E(z)$, and potential, $V(z)$, are related by

$$E(z) = -\frac{\partial V(z)}{\partial z}. \tag{3.54}$$

Hence, for a bias of V_{a} across a heterostructure of length L, the applied electric field is[9]

$$E_{\mathrm{a}} = -\frac{V_{\mathrm{a}}}{L}, \tag{3.55}$$

[8] In practice, $E_{k_{\|}}$ is calculated for a number of values of $k_{\|}$, and a cubic spline fit is used to approximate the full curve [67].

[9] Assuming the same dielectric constant for each material in the region over which the field is applied.

or, alternatively, $V_a = -E_a L$. Therefore, the electric potential within the heterostructure is given by

$$V_a(z) = -E_a z. \tag{3.56}$$

The Hamiltonian of the perturbed system, $\widehat{H}'$, is obtained by adding the electron energy due to the applied field to the unperturbed Hamiltonian, $\widehat{H}$,

$$\widehat{H}' = \widehat{H} + (-|e|)V_a(z) \tag{3.57}$$

$$= \widehat{H} + |e|E_a z. \tag{3.58}$$

From (3.58), one can see that the effect of a uniform positive field is a linear increase in the electron's energy with increasing z, i.e., the electronic energy bands are no longer 'flat'. This can be approximated by a series of small flat steps, as shown in Figure 3.4. Within the region of each step the Schrödinger equation consists of the original Hamiltonian plus a constant potential energy term. Thus, a heterostructure in an applied electric field can be treated as before, but with each material layer being sub-divided into smaller layers, which have their band edges offset against each other in the same way as different materials, but according to the stepped values of $V_a(z)$.

Clearly, the steps should be as narrow as possible to maximise the accuracy of the approximation. However, since the EPM includes the atoms explicitly, there is a lower limit on this width, namely the width of a monolayer. A monolayer consists of two atomic layers of each atom in the zinc-blende material, i.e., a layer of each of the atoms in the primitive unit cell.

Also, in these calculations, we make the additional approximation that the electric field is applied only over the heterostructure's 'well' region, i.e., the extreme left

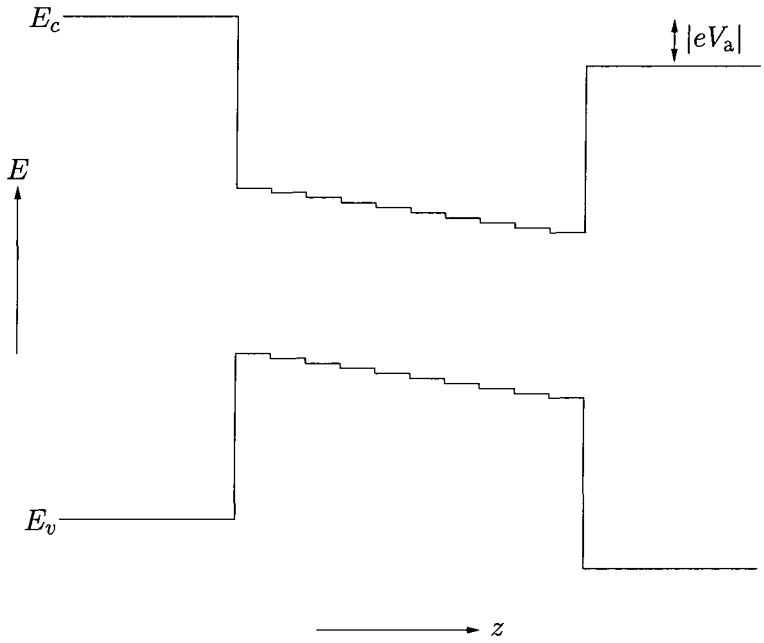


Figure 3.4 Schematic illustrating the piece-wise flat approximation used in the calculation of heterostructure bound states in an applied electric field. In the present work, the electric field is not applied to the extreme left and right barriers, allowing for genuinely bound states. The figure shows the effect of the electric field on the electron energy band edges, hence, since electrons carry negative charge, this particular figure corresponds to a negative applied electric field.

and right barriers both remain completely flat, with the right barrier being offset by $-|e|V_{\mathrm{a}}$ relative to the left barrier. The monolayers of the materials in between the barriers, the 'well' region, then have their bands offset accordingly. Thus, the l^{th} monolayer is band offset (relative to its original offset, if any) by

$$V_l = -|e|V_{\mathrm{a}}(z_l), \tag{3.59}$$

where z_l is the position of the middle of the monolayer, with respect to the left barrier. Since bound electrons tend to concentrate in the well region, the effect of the flat barriers is minimal, at least for small fields.

Chapter 4

Heterostructures — results

4.1 Introduction

The purpose of this chapter is to demonstrate the power of the chosen EPM complex band structure method for performing heterostructure calculations, whilst investigating two types of heterostructure based on the InAs/GaSb/AlSb materials system. The heterostructures studied are far more demanding of the techniques used in the analysis than the more common GaAs/AlGaAs systems. Because of this, these materials result in heterostructures that are also physically more interesting than typical systems, and although they have received a fair amount of attention from other workers (see later references), these particular material systems are still quite new and hence have been investigated on a much smaller scale than more established materials.

In Section 4.2 we investigate the electronic properties of the InAs conduction band well formed between AlSb barriers. This forms a deep type II well, and so attention is naturally focused on transitions within the conduction band, although the system has other applications which do not involve excited states, such as InAs/AlSb superlattice Schottky diodes [69]. Other workers have investigated

some aspects of this system using alternative theoretical approaches, specifically tight-binding [10], $\mathbf{k} \cdot \mathbf{p}$ [70,71] and *ab initio* pseudopotential [19] methods, as well as experimentally [6,69,71]. Here, we use the empirical pseudopotential method to investigate the possibilities of transitions corresponding to the technologically important 1.55μm electromagnetic wavelength. Since the AlSb barriers are indirect gap materials, with the conduction band X-valley considerably lower in energy than the Γ-valley, a realistic description of the conduction band throughout the Brillouin zone is important. This favours the use of EPM based methods for calculating properties associated with intraband transitions.

Section 4.3 investigates the asymmetric, 'stepped' well formed by layers of GaSb and InAs sandwiched between AlSb barriers. Heterostructures consisting of these three materials are arousing increasing interest, chiefly because of their unusual energy band line-ups and the wide flexibilty in device design they offer. Indeed, in their brief review paper, *Prospects for the future of narrow bandgap materials* [72], McGill and Collins describe ten different electron or hole tunneling devices achievable purely by adjusting the growth sequence of InAs, GaSb and AlSb.

One aspect of this system that is of particular interest is the 'band overlap' of InAs and GaSb, where the InAs conduction band minimum lies below the GaSb valence band maximum [73]. This has led other workers [69,74–76] to investigate the possibilities of resonant interband tunneling (RIT) diodes with large negative differential resistance (NDR). Another application of this material system is the ability to design and fabricate normal incidence infrared electroabsorption modulators [77–79], where the absorption is via *intra*band transitions in the L-satellite valleys. Other applications discussed in the literature include infrared lasers based on intersubband transitions [80], large non-linear responses with third harmonic generation [81] and lasers/detectors employing a combination of intersubband transitions and both intra and interband tunneling [82].

Material	Γ_v	Γ_c	X_c	L_c
AlSb	−0.189	2.128	1.425	2.008
InAs	−0.299	0.000	1.793	1.096

Table 4.1 The important energy band line-ups of the InAs/AlSb system, as calculated by the EPM discussed in the text. The materials are in-plane stressed to lattice match GaSb. All the energies are with respect to the InAs conduction band minimum, and are in eV. References are given in Appendix D.

In this work, we analyse how the stepped well bound state energies are effected by variations in the InAs layer width and by the application of an external static electric field.

4.2 Single quantum wells — InAs/AlSb

We first turn our attention to the analysis of a single InAs/AlSb quantum well. These materials form a type II staggered heterojunction, with the InAs valence band maximum below that of the AlSb. Thus, due to the small band gap of InAs and the relatively large gap of AlSb, an unusually deep well is formed in the conduction band, see Table 4.1 and Figure 4.1. In fact, AlSb has an indirect band gap, so that truly bound states can only exist up to the X-valley minimum, rather than the considerably higher Γ-valley minimum, but this still produces a very deep well.

Since the two materials are usually grown on a GaSb substrate or buffer layer, in this work they are strained in-plane to match the bulk lattice constant of GaSb. The AlSb is in biaxial compression while the InAs is in biaxial tension, corresponding to in-plane strains of approximately −0.67% and +0.61%, respectively. The materials are relaxed elastically in the growth direction (taken to be the z-direction), as described in Section 2.3.1. The strain is also included for consistency because the heterostructure considered in Section 4.3 includes a GaSb layer, and using a common in-plane lattice parameter allows direct comparison between the various

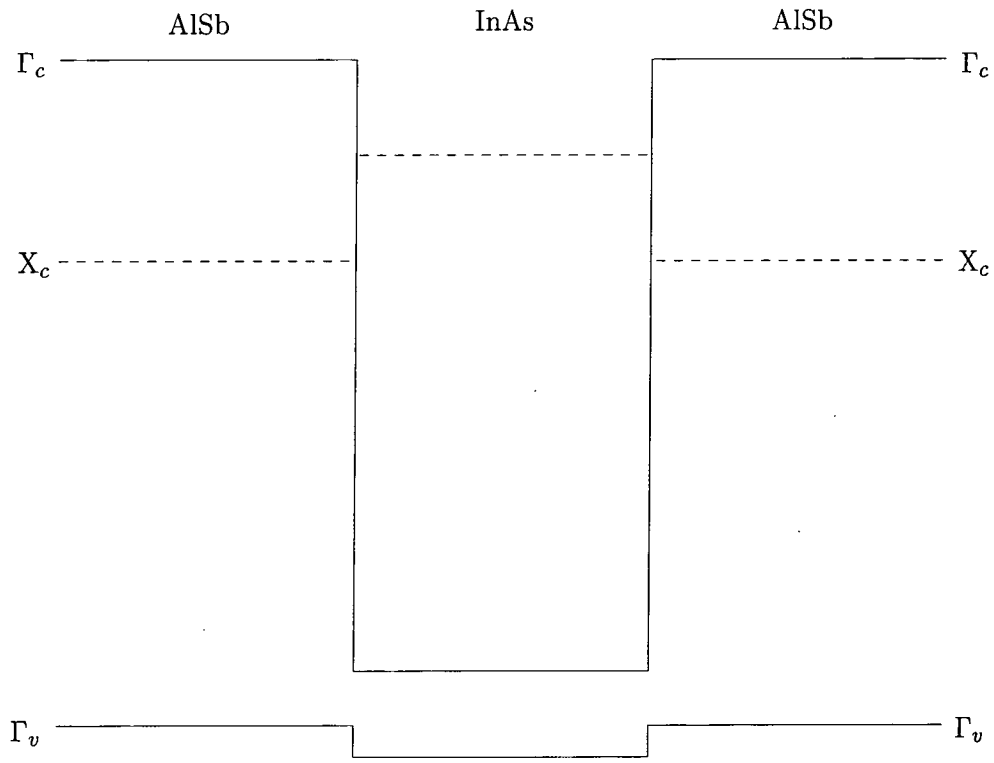


Figure 4.1 InAs/AlSb energy band line-ups. The bands are drawn to scale, with the numeric values given in Table 4.1.

structures. In any case, the lattice mismatch between the three materials is small (about $\pm 0.6\%$) and thus including strain does not have a major effect on the conduction band, unlike the valence band [37].

From Table 4.1, the well depth is 1.425eV, which is much greater than that of the more commonly used quantum well systems, such as $InP/In_{0.53}Ga_{0.47}As$ or GaAs/AlGaAs, which have conduction band well depths in the range 0–0.3 eV. The large depth allows for the possibility of transitions between bound states within the well (intraband transitions) at energies around 0.8 eV, corresponding to the radiation wave length of 1.55μm, which is used in fibre optic telecommunications (see, for example, p104 of [83]). This section concentrates on this transition. The pseudopotential calculations of the quantum well energy band structure and momentum matrix elements use 89 reciprocal lattice vectors and the first 9 in-plane projections, $\mathbf{g}_{||}$, in the wavefunction matching procedures (see Section 3.3).

4.2.1 Energy vs well width

The variation in the bound state energies with well width is shown in Figure 4.2. This also shows the energy variation predicted by a simple 1D single parabolic band (SPB) approximation (see Appendix C). To ensure a meaningful comparison between the two methods, the parameters for the SPB approximation are derived from the EPM band structure. From the figure, we see that the SPB prediction for the ground state energy is in excellent agreement with the results of the full EPM calculation, but the excited states disagree dramatically. This is because the EPM excited states have a sizeable contribution from the complex-k branch leading down from the X-valley (see Figure D.10), which is totally absent from the SPB. Thus, the SPB approximation is wholly inadequate for the study of properties involving any excited states and we expect the same to be true for 8 band $\mathbf{k} \cdot \mathbf{p}$ methods.

The figure shows that there is a minimum width of 7 monolayers for a bound excited state to exist. In particular, at a well width of 7 monolayers there is a single excited state at an energy 0.816 eV above the ground state. Although this is quite a narrow well ($7 \times a_{\perp}^{(\mathrm{InAs})}/2 \approx 21$Å), its fabrication is certainly possible [6].

Before considering the properties of the 7 monolayer wide well, we compare the bound state energies predicted by the present theory with those reported by other workers. Bound state energy data for this system is somewhat limited, which is of course one of the motivations for its study.

Boykin [10] uses a nearest-neighbour 10 band tight-binding model to calculate the energies of quasi-bound states in an InAs well between two AlSb barriers. Table 4.2 compares the ground state energy, as a function of well width, as predicted in the present work and by Boykin. The predictions are in reasonable agreement, but it is impossible to make an absolutely meaningful comparison since many details of Boykin's calculation are unknown, such as the AlSb energy band edges and the

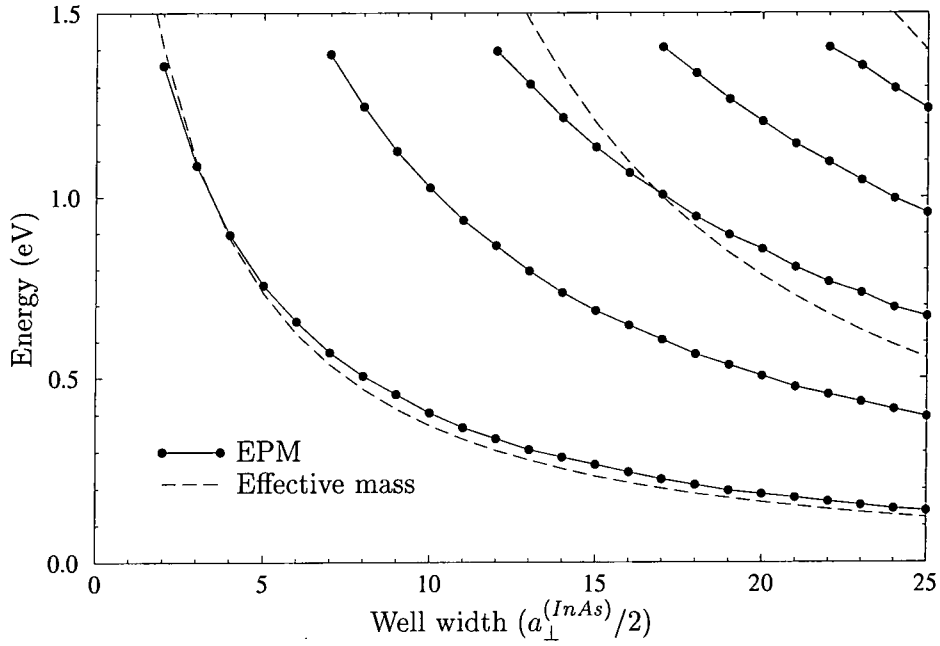


Figure 4.2 Bound state energies as a function of InAs well width, with AlSb barriers. The materials are in-plane strained to match the bulk lattice parameter of GaSb. The energy is with respect to the bottom of the well, which is 1.425 eV deep. The dashed lines show the results of a 1d parabolic band calculation, using data derived from the empirical pseudopotential method (EPM).

	Energy (eV)	
InAs width ($a/2$)	This work	Boykin
2	1.36	1.2
3	1.09	0.95
4	0.90	0.8
5	0.76	0.7

Table 4.2 Comparison of the ground state energy predicted in the present work and by Boykin [10] for a narrow InAs/AlSb quantum well of various widths. The energy origin is at the InAs conduction band minimum.

temperature at which the modelling is being done. However, it is not surprising that Boykin reports slightly lower energies, since in general the effect of the finite barrier widths is to slightly reduce the quantum confinement of the quasi-bound states, and hence their energy, with respect to the genuinely bound states of this work.

Warburton *et al.* [70] report both theoretical and experimental results for a 150Å wide well. Table 4.3 makes a comparison between their theoretical predictions and those of this work, where we have used a well width of $50a/2$. They also report an experimentally measured ground to first excited state (1–2) transition energy of 0.134eV. As was the case with Boykin's data, direct comparisons are difficult, especially since Warburton *et al.* report data obtained from low temperature experiments and modelling, whereas this work models the materials at room temperature. Besides, comparison between theoretical and experimental data is often awkward as there is always uncertainty over the exact experimental parameters and measurements, and for purely practical reasons some effects must be omitted from the theoretical calculations. However, in this case the results are still in reasonable agreement.

Experimental and theoretical predictions for the ground state energy have also been reported by Fuchs *et al.* [71], though unfortunately, these are also low temperature results. They report interband, spatially indirect transition energies, i.e.,

	Energy (eV)	
State/Transition	This work	Warburton *et al.*
1	0.05	0.06
2	0.16	0.18
3	0.29	0.33
1–2	0.11	0.12
1–3	0.24	0.27
2–3	0.13	0.15

Table 4.3 Comparison of the energies of conduction band bound states, and transitions between them, as predicted in the present work and by Warburton *et al.* [70] for a 150Å wide InAs/AlSb quantum well. The energy origin is at the InAs conduction band minimum.

those corresponding to transitions from the top of the AlSb valence band to the InAs quantum well ground state. For a 7.5nm ($25a/2$) quantum well, in-plane matched to GaSb (as is the case for the present work), they predict and measure a transition energy of approximately 0.44eV, in contrast to the value of 0.33eV predicted in the present work.[1] At first glance, this seems to cast the results presented in this work in a considerably less favourable light than the comparisons with Boykin and Warburton *et al.* However, unlike those previous comparisons, this data directly includes the InAs principle band gap. Since the gap is relatively small, a small change in its value due to temperature variations is more noticeable than if the band gap was larger. If we take into account the increased band gap of the low temperature InAs (Fuchs *et al.* assume a value of 0.357eV, whereas we are using the room temperature, strained value of 0.299eV), then the prediction of this work becomes 0.39eV (assuming no other changes). The difference between this value and that reported by Fuchs *et al.* is then of the same order as for the comparisons with Boykin and Warburton *et al.*

[1] From Table 4.1, we note that the data in Figure 4.2 must be increased by 0.189eV to obtain the energy of the well states with respect to the top of the AlSb valence band.

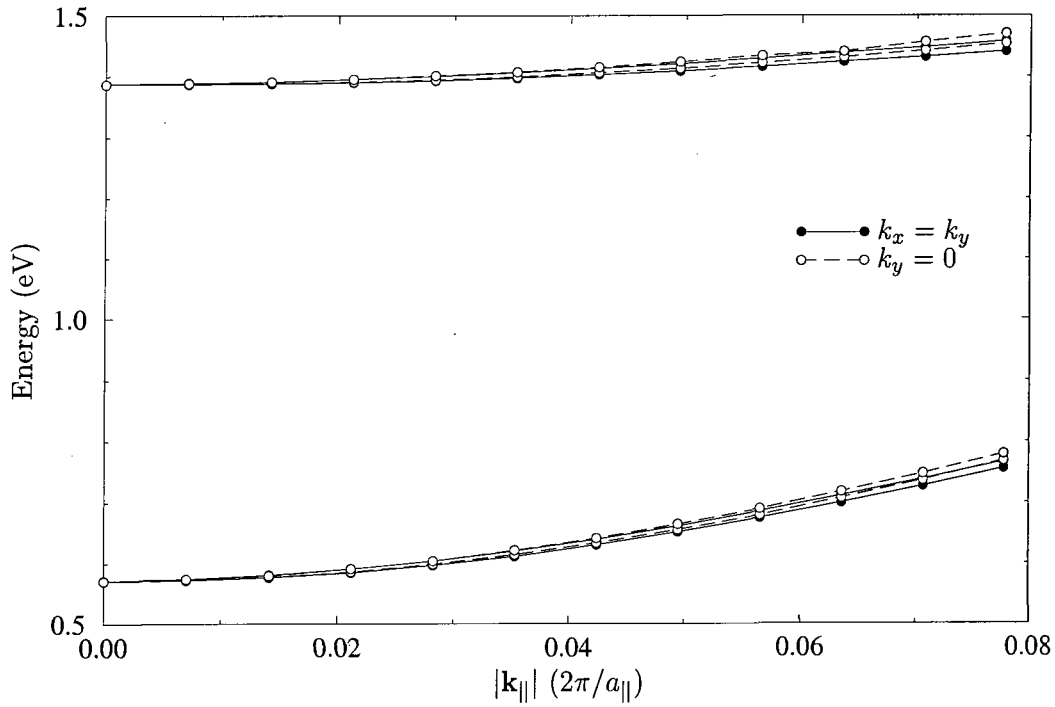


Figure 4.3 In-plane energy dispersion of the two conduction band bound states in the 7 monolayer InAs/AlSb SQW, as a function of the magnitude of the in-plane wave vector, $|\mathbf{k}_{||}| = |(k_x, k_y)|$. The close agreement between the curves for $k_x = k_y$ and $k_y = 0$ indicate an almost isotropic energy dispersion in $|\mathbf{k}_{||}|$. The energy scale is with respect to the bottom of the well, i.e., the bottom of the InAs conduction band.

4.2.2 Energy vs $\mathbf{k}_{||}$

The variation of energy with in-plane wave vector, $\mathbf{k}_{||}$, in the two bound states of the 7 monolayer system is plotted in Figure 4.3. The energy dispersion is plotted as a function of the magnitude of the in-plane wave vector for both $\mathbf{k}_{||} = (k, 0)$ and $\mathbf{k}_{||} = (k, k)$. It is apparent that the in-plane energy band structure is very nearly isotropic, which is to be expected because the states are dominated by the s-type component associated with the bulk InAs conduction band Γ-valley. One can also see the lifting of the Kramers degeneracy away from the high symmetry point at $\mathbf{k}_{||} = (0, 0)$. However, the splitting is very small in this case, being only about 10 meV, so we can reasonably approximate both ground and excited states as spin degenerate for non-zero $\mathbf{k}_{||}$, using the mean energy of the split bands.

Another feature of the ground and first excited state dispersion curves is that they

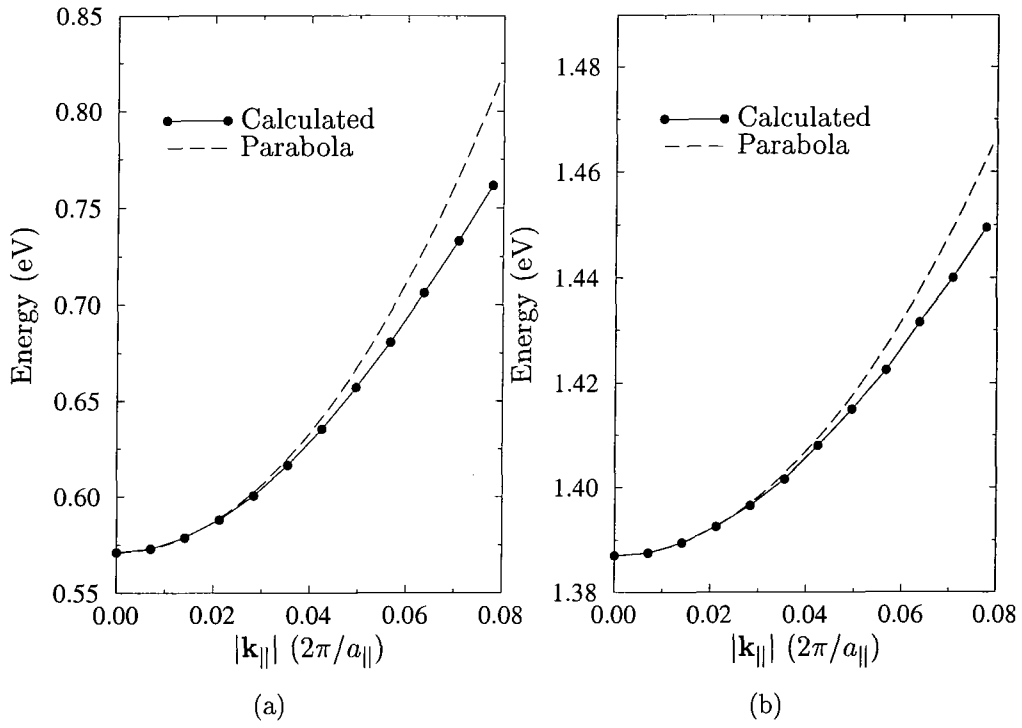


Figure 4.4 Demonstration of the non-parabolicity of the in-plane energy dispersions for the 7 monolayer InAs/AlSb SQW. (a) Ground state (b) Excited state. Both energy scales are with respect to the bottom of the InAs well.

are not parallel, and indeed, their energy difference is reduced with increasing in-plane wave vector. This is due to the marked non-parabolicity of the InAs conduction band Γ-valley. Since the conduction band varies sub-quadratically even a small distance away from the Γ point, the effective mass increases and causes the in-plane dipersion for the excited state to have a lower curvature than that of the ground state [84].

The non-parabolicity of the in-plane subband is illustrated in Figure 4.4. Although the subband energies vary quadratically for very small $|\mathbf{k}_{||}|$ they become noticeably sub-parabolic at magnitudes well before the largest value shown in Figure 4.4, $|\mathbf{k}_{||}| = 0.08(2\pi/a_{||})$.

Since we are interested in transitions between the two states, it is useful to plot their energy separation as a function of the magnitude of the in-plane wave vector

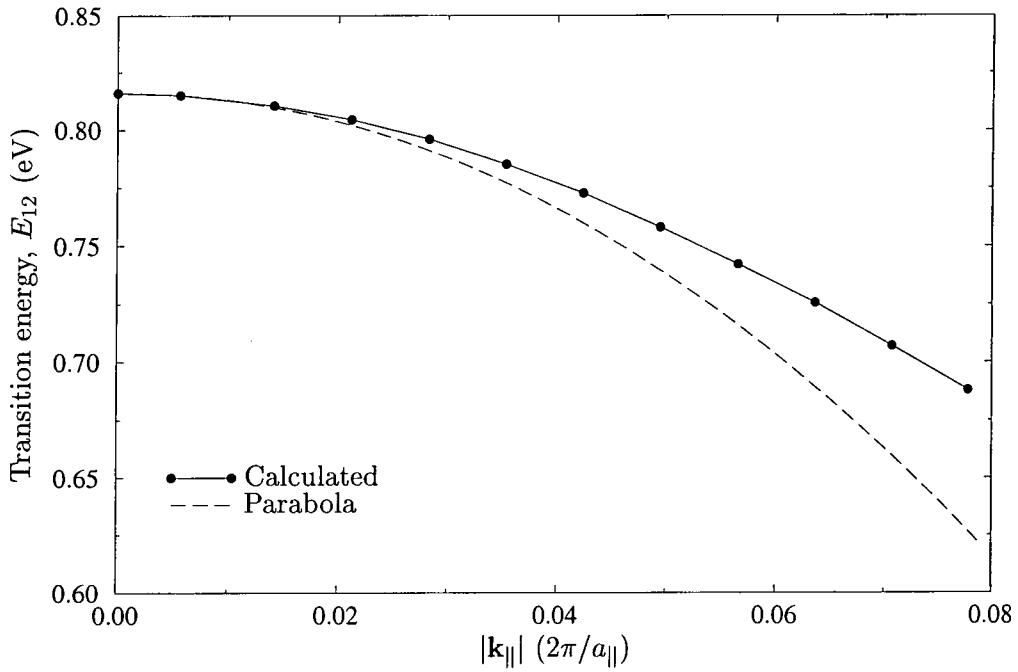


Figure 4.5 The vertical transition energy between the ground and excited states of the 7 monolayer InAs/AlSb SQW, as a function of the in-plane wave vector magnitude. A parabola is also shown to demonstrate the non-parabolicty.

— this is shown in Figure 4.5. In fact, Figure 4.5 shows the energy difference between the mean values of the ground and excited state, i.e., the lifting of the Kramers degeneracy is neglected, since it is so small. The plot shows that although the transition energy does vary with in-plane wave vector, it is within 5% of 0.8 eV for wave vector magnitudes up to about $0.05(2\pi/a_{||})$. As a rough guide to how satisfactory this is for the possibility of (room temperature) devices, we consider the (classical) thermal energy, k_bT, of a 2D particle. Equating this to an effective mass expression for the transition energy (with the effective mass, m^*_{12}, fitted to the curvature in Figure 4.5), $\hbar^2k^2_{||}/2m^*_{12}$, leads to a thermal energy of 26meV (at 300K) and a corresponding wave vector magnitude of $0.025(2\pi/a_{||})$. This naive picture suggests that the changes in E_{12} with $k_{||}$ would not hinder transitions at around 0.8eV.

4.2.3 Joint density of states

The joint density of states (see Section 3.7) between the conduction band ground and first excited bound states of the 7 monolayer system is shown in Figure 4.6. The non-parabolicity alters the joint density of states from the simple 'text book' form. If parabolic subbands are assumed then the transition energy between bands i and j, with effective masses m_i^* and m_j^*, is

$$E_{k_{||}} = \frac{\hbar^2 k_{||}^2}{2}\left(\frac{1}{m_j^*} - \frac{1}{m_i^*}\right). \tag{4.1}$$

Therefore, from (3.53), the joint density of states is given by

$$g_{ij}(E) = \hbar^{-2}\pi^{-1}\left|\frac{m_i^* m_j^*}{m_i^* - m_j^*}\right|. \tag{4.2}$$

Thus, using the parabolic subband approximation, we expect the JDOS to be either constant or a delta function, for differing or equal subband effective masses, respectively. Although the two subbands are not quite parallel, the joint density of states is still increased compared to interband transitions, since the curvature of the two subbands has the same sign for intraband transisitions in the conduction band.

4.2.4 Momentum matrix elements

The optical transition rate between states is proportional to the squared magnitude of the momentum matrix elements between them. The component of the momentum operator of interest in calculating the momentum matrix element depends on the polarization of the electromagnetic (EM) radiation. The two cases of particular interest are:

- An EM plane wave propagating in the growth direction, resulting in an electric field component in an in-plane direction. This is termed the *Transverse*

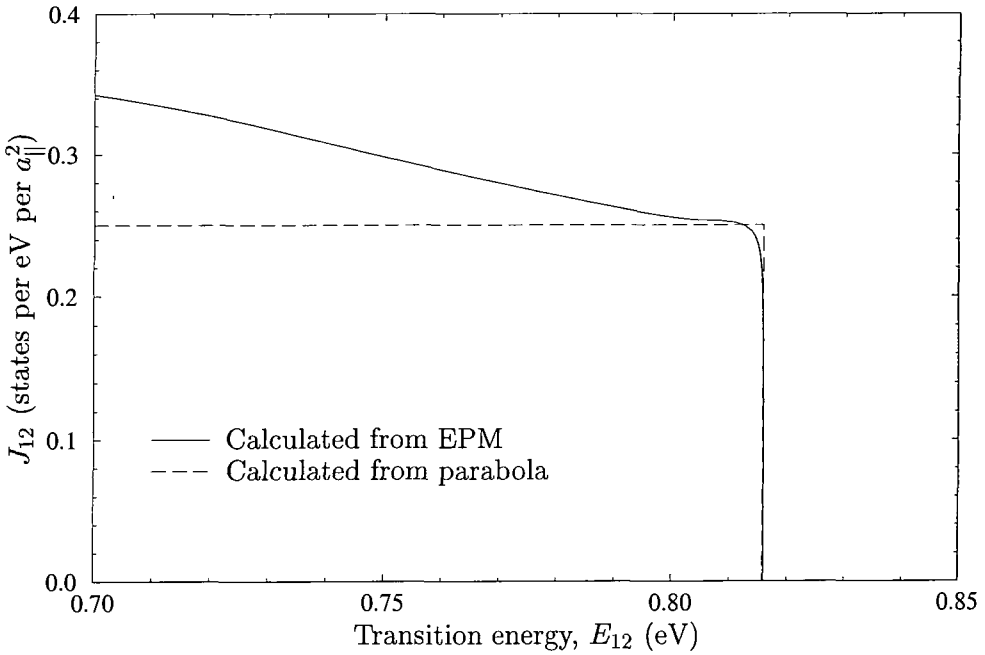


Figure 4.6 Joint density of states in energy, J_{12}, between the ground and excited conduction band states of the 7 monolayer InAs/AlSb SQW. This illustrates how the non-parabolicity of the EPM in-plane band structure results in deviations from the simple stepped form predicted by assuming parabolic energy bands.

Electric (TE) mode.

- An EM plane wave propagating in an in-plane direction, giving an electric field component in the growth direction. This is termed the *Transverse Magnetic* (TM) mode.

Of course, there are still two degrees of freedom in the in-plane direction. Thus, the transition rate due to the TE mode depends on the specific polarization of the electric field component in the plane of the well and the orientation, as well as the magnitude, of the in-plane electron wave vector.

The squared magnitude of the momentum matrix elements between the ground and excited states of the 7 monolayer InAs/AlSb system are plotted in Figure 4.7 as a function of $|\mathbf{k}_{\|}|$. Curves for both the TE and TM modes are shown, but the TE squared matrix element has been scaled up by a factor of 10 for clarity. The key

features are the slow drop in squared matrix elements for both modes with increasing in-plane wave vector magnitude, and the disparity in the values for the two modes, in agreement with the $\mathbf{k} \cdot \mathbf{p}$ predictions of Warburton *et al.* [70]. In fact, at $|\mathbf{k}_{||}| = 0$, the squared matrix element for the TE mode is only approximately 0.5% of that for the TM mode. The difference in parity of the wavefunctions of the two states in the growth direction results in a large dipole matrix element between them, along that direction, and hence we also obtain a large momentum matrix element in the growth direction. This situation arises in the TM mode. In the plane of the well we have bulk-like states dominated by contributions from the same band (conduction, in this case) with no parity difference between the wavefunctions, hence the very small matrix elements for the TE mode.

Simple effective mass and envelope function theories predict that the matrix elements for the TE mode vanish. The small but non-zero matrix elements arise here because of the slight band mixing of the bound states. The matrix elements have been demonstrated experimentally to be non-vanishing by Liu *et al.* [85] for GaAs based quantum wells. They obtained TE:TM signal ratios of 0.2% for a 54Å wide $GaAs/Al_{0.26}Ga_{0.74}As$ quantum well and 3% for a 37Å wide $In_{0.1}Ga_{0.9}As/Al_{0.44}Ga_{0.56}As$ quantum well. As they note, there is still some debate over the correct order of magnitude of TE:TM ratios, with some $\mathbf{k} \cdot \mathbf{p}$ calculations predicting ratios as high as 10%. A recent theoretical study by Batty and Shore [86] predicts a TE:TM ratio of no more than 1% for various systems. Thus, the author believes the prediction here of a ratio of 0.5% to be at least qualitatively correct.

4.2.5 Applied electric field

We now turn our attention to the response of the system to a static electric field, E_a, applied in the well growth direction. We are interested in how the bound state energies change with the electric field.

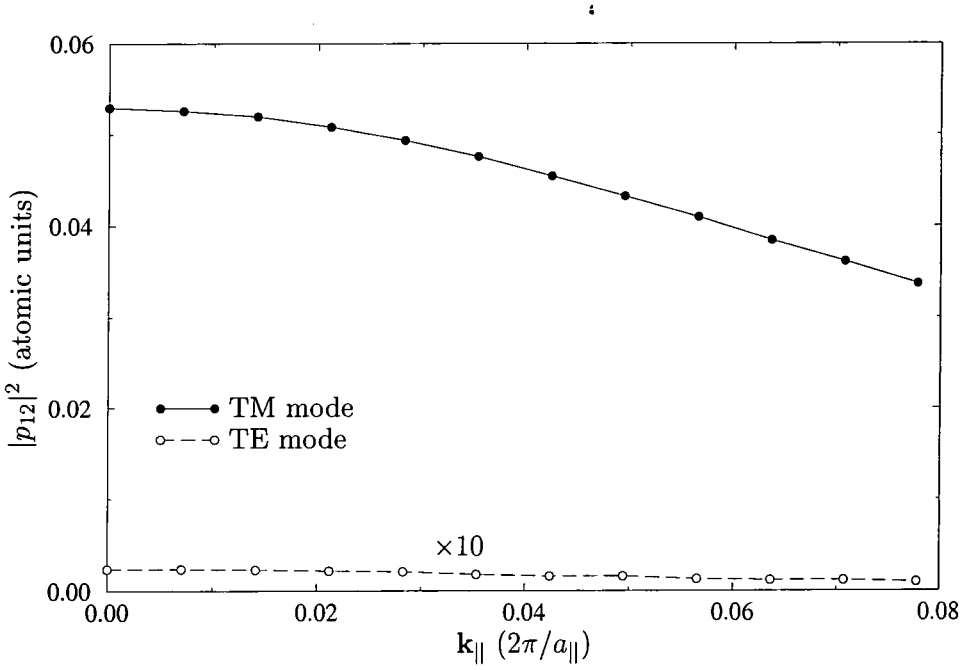


Figure 4.7 Squared magnitude of the momentum matrix elements between the ground and excited conduction band states of the 7 monolayer InAs/AlSb SQW. The plot for the TE mode has been magnified by a factor of 10.

The application of an electric field distorts the band profile of a heterostructure, so that the usual energy reference point, the top of the valence band, becomes a function of position along the well growth direction. Therefore, care must be taken when choosing the energy origin so that the shift in energy of the band edges (linear with the applied field) does not obscure the important physics. Therefore, we choose the origin of the applied electrostatic field to be midway between the left and right barriers. Chosen in such a way, the average of the electrostatic potential energy is zero, if the bound state energies are measured with respect to one of the energy bands (conduction, in this case) at the centre of the well, i.e., if the applied field origin spatially coincides with the energy reference point, then that reference point is unchanged. Equivalently, we calculate the change in the difference between the bound state energy and the average band edge potential energy in the well, as a function of applied field. Of course, the only experimentally measurable energies are

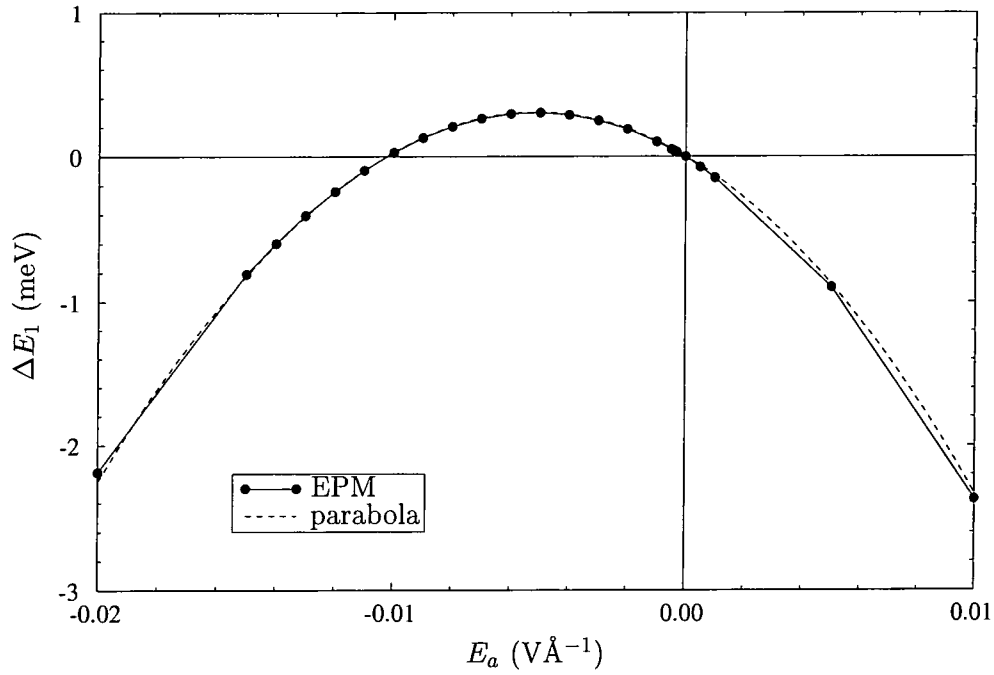


Figure 4.8 The change in energy with applied perpendicular electric field (Stark shift) of the ground state in a 7 monolayer InAs/AlSb single quantum well. The offset of the maximum from the origin is due to a lack of inversion symmetry in the potential of the well. The in-plane wave vector is zero.

those corresponding to transitions between states, in which case the energy reference point is naturally irrelevant.

Figure 4.8 shows the change in the ground state energy with applied electric field of the 7 monolayer InAs/AlSb quantum well. For applied fields with a 'small' magnitude (i.e., those that correspond to applied biases that are small compared to the well depth), envelope function theories would predict a quadratic energy dependence on the field because of the reflection symmetry of the quantum well structure [18,87]. This is not seen in Figure 4.8, where the energy maximum does *not* coincide with zero applied field but is displaced slightly. This is a consequence of the atomic resolution of the empirical pseudopotential method employed here. The quantum wells considered in this chapter are all an integer number of monolayers wide, with a monolayer being half the lattice parameter. A single monolayer contains two (001) planes of the atoms that make up the binary compound, each atomic layer

containing only atoms of the same type. So, for a 2 monolayer InAs/AlSb conduction band well, we have atomic layers arranged in the form

$$\cdots \mathrm{Sb} \cdot \mathrm{Al} \cdot \mathrm{Sb} \cdot \mathrm{Al} \,\Big|\, \mathrm{As} \cdot \mathrm{In} \cdot \mathrm{As} \cdot \mathrm{In} \,\Big|\, \mathrm{Sb} \cdot \mathrm{Al} \cdot \mathrm{Sb} \cdot \mathrm{Al} \cdots$$

It is apparent from the atomic arrangement that on the microscopic scale, the potential will not be symmetric about the centre of the well.

The lack of symmetry in the microscopic potential results in a corresponding lack of symmetry in the bound state wavefunctions and hence a small dipole moment exists at zero bias. Therefore, when applying an electric field, either the positive or negative bias will expend energy to overcome the small opposite built-in dipole moment of the wavefunction, before inducing the larger dipole moment dependent on the field. This feature can be reproduced using simple envelope function models (such as the single parabolic band approximation) by introducing a small step in the potential.

It should be noted that this offset in the maximum of the response curve only occurs because of the perfect nature of the crystals when modelled by the EPM and the limitation to integer monolayers of the well width. Including an extra atomic layer in the well region would result in the usual response with the maximum at zero bias (cf. the symmetric charge densities thus obtained in Chapter 6). Apart from the displacement of the energy maximum, Figure 4.8 shows that the change in the ground state energy has the parabolic response to the applied field, as predicted by envelope function methods.

Of more physical significance is how the energy difference between the ground and excited states varies with the application of the electric field. This is important because devices based on heterostructures need a bias to be applied so that the carriers involved in the desired optical transitions can be introduced or extracted

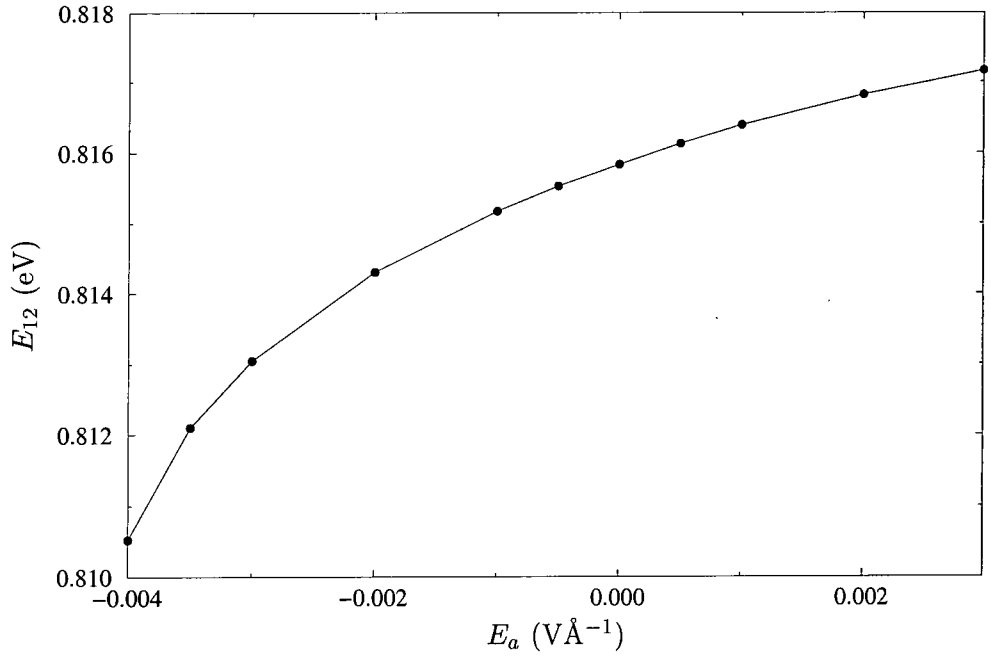


Figure 4.9 The variation with applied electric field of the transition energy, E_{12}, between the ground and first excited states of the 7 monolayer wide InAs/AlSb single quantum well.

to the well region, as required by the particular functioning of the device. For example, a photodetector using the intraband transition discussed in this section would require a bias to alter the barrier potential profile to allow the electrons in the bound excited state to escape the well by tunnelling through the barriers.

The variation in the intersubband energy with applied electric field is plotted in Figure 4.9. The change in the transition energy is small over the range of fields typically used in devices, which suggests that this system might be a viable candidate for photodetection at 1.55μm.

4.3 Asymmetric 'stepped' quantum well

Having examined the case of the relatively simple, albeit unusually deep, single quantum well, we now move on to a more complicated heterostructure. Specifically,

Material	Γ_v	Γ_c	X_c	L_c
AlSb	−0.189	2.128	1.425	2.008
GaSb	0.211	0.937	1.257	1.101
InAs	−0.299	0.000	1.793	1.096

Table 4.4 The important energy band line-ups of the InAs/GaSb/AlSb system, as calculated by the EPM. The materials are in-plane stressed to lattice match GaSb. All the energies are with respect to the InAs conduction band minimum, and are in eV. References are given in Appendix D.

we examine an asymmetric[2] well formed by sandwiching layers of both InAs and GaSb between the AlSb barrier layers. The resulting band line-up is illustrated in Figure 4.10, with the numerical values listed in Table 4.4. The band edge line-up shows that this system is particularly interesting because not only does it form an asymmetric conduction band well but also, due to the type II broken gap heterojunction formed between the InAs and GaSb layers, the InAs part of the conduction band well overlaps in energy with the valence band well formed in the GaSb layer.

To understand the results, it helps to split the energy range over which bound states may exist into three regions, as labelled in Figure 4.10. We can then characterise the states in each energy region as follows (following the labelling of the figure):

1. Both InAs and GaSb conduction band states are allowed. In region (a) there is also a well formed in the X-valley of GaSb so that there will be bound X-states present. In region (b) only Γ-like states are present in both the GaSb and InAs layers.

2. This region corresponds to the GaSb band gap, so only Γ-like InAs conduction band bound states are present.

[2]The term 'asymmetric' is used here to describe a macroscopic asymmetry in the heterostructure, due to the choice in the materials forming the structure, rather than the small asymmetry in the potential on the atomic scale discussed in Section 4.2.5.

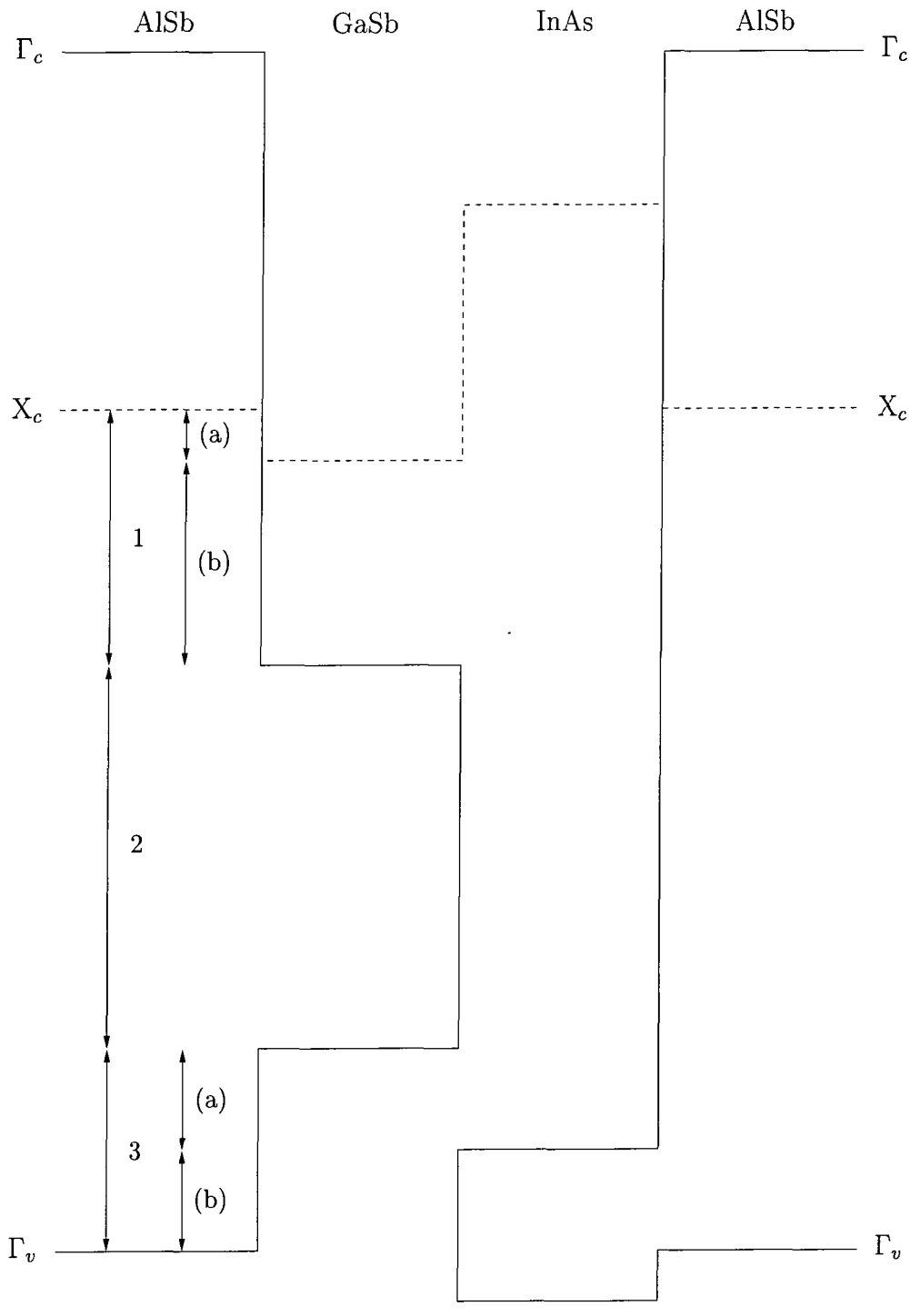


Figure 4.10 Band profile for the InAs/GaSb/AlSb asymmetric, or 'stepped', quantum well, showing the relative positions of the zone centre and X-valley band edges. The values for the labelled energy ranges, in which bound states can exist, are given in Table 4.4.

3. Valence band GaSb states are present. In region (a) there may also be InAs conduction band states, whilst region (b) includes the InAs band gap, so only GaSb states are allowed. Region (a) is the band overlap region.

Of course, quantum tunneling means that in general a bound state will be spatially spread over each material region to a certain extent.

4.3.1 Energy vs InAs layer width

Figure 4.11 shows the variation in energy with InAs layer width of bound states with energies around the band overlap (labelled 3(a) in Figure 4.10), for a fixed GaSb layer width of 20 monolayers ($a_\perp/2$). For narrow InAs layers, the conduction band bound states lie well above the top of the GaSb valence band and so the valence and conduction band wells act as two separate single wells. As the InAs layer width increases, the conduction band states drop in energy so that, at an InAs width of $20a_\perp/2$, the conduction band ground state energy drops below the GaSb valence band maximum. Increasing the InAs width further causes the conduction and heavy hole ground states to cross. However, due to the coupling of the conduction and light hole states, the conduction and light hole ground state bands actually anti-cross.

This anti-crossing behaviour is illustrated more clearly by Figure 4.12, which shows a similar plot to that in Figure 4.11 but with the bound state energies of a single 20 monolayer wide InAs/AlSb conduction band quantum well overlaid. This figure also shows more clearly the smaller amount of anti-crossing between the electron ground and heavy hole first excited states, resulting from the small amount of band mixing in the two states. In addition, the figure demonstrates how the light hole ground state begins to anti-cross with the excited conduction band state as the InAs width is increased further.

The anti-crossing behaviour is analogous to the opening of so-called 'minigaps'

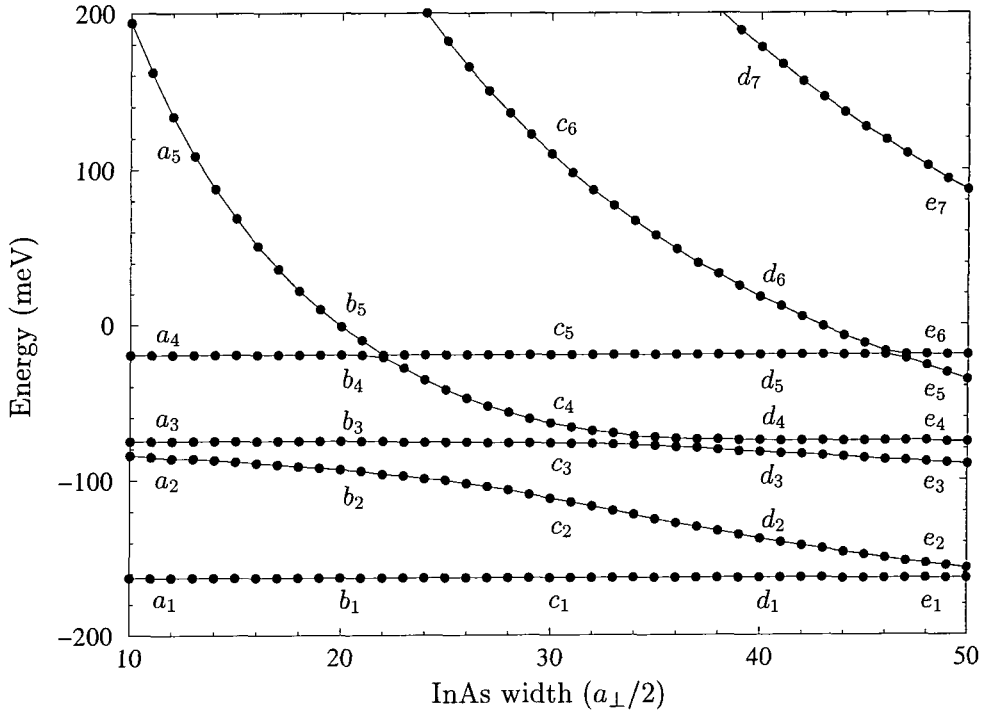


Figure 4.11 Variation in energy of bound states around the band overlap of the InAs/GaSb/AlSb stepped quantum well, for a $20a_\perp/2$ wide GaSb layer and zero in-plane wave vector. The energy origin is at the top of the GaSb valence band, with the bottom of the InAs conduction band being at −211meV. The bands are labelled for reference in the text.

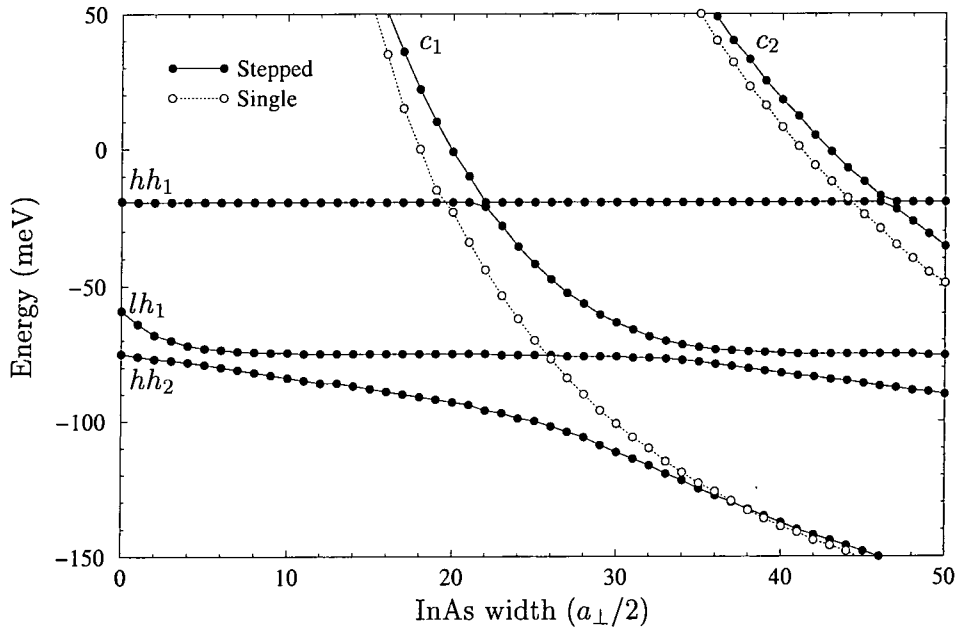


Figure 4.12 Comparison of the bound state energies in the InAs/GaSb/AlSb and InAs/AlSb quantum well systems, for zero in-plane wave vector. The GaSb layer is 20 monolayers wide, as in Figure 4.11 (note that in this figure the InAs width starts at zero and the energy scale has been magnified slightly). The bands are labelled according to their nature for zero InAs layer width. The energy origin is at the top of the GaSb valence band, with the bottom of the InAs conduction band being at −211meV (note that Figure 4.2 has the energy origin at the top of the InAs conduction band).

[88–90] in the in-plane band structure, which occurs when the widths of the spatially adjacent (or very close) InAs and GaSb layers are such that one or more electron states is lower in energy than at least one hole state at the 2D Brillouin zone centre, $\mathbf{k}_{||} = (0, 0)$. This has been attributed to hybridization of the electron and hole states when they are close in energy. When hybridization occurs, the wavefunctions of the two states have both electron and hole-like characteristics in the InAs and GaSb layers, respectively, and the associated probability density is distributed in both material layers in approximately equal measure [88]. However, the situation described in Figures 4.11 and 4.12 is slightly different, as is explained below.

Figure 4.13 shows the in-plane averaged probability densities of the electron, heavy and light hole ground states, plus the first excited heavy hole state, at InAs layer widths of 20, 30 and $40a_{\perp}/2$. The labelling of the states is consistent with that of Figure 4.11. When the InAs width is $20a_{\perp}/2$ (Figure 4.13(a)), we can clearly assign the bands to states as follows:

$$b_2 \mapsto lh_1$$

$$b_3 \mapsto hh_2$$

$$b_4 \mapsto hh_1$$

$$b_5 \mapsto c_1.$$

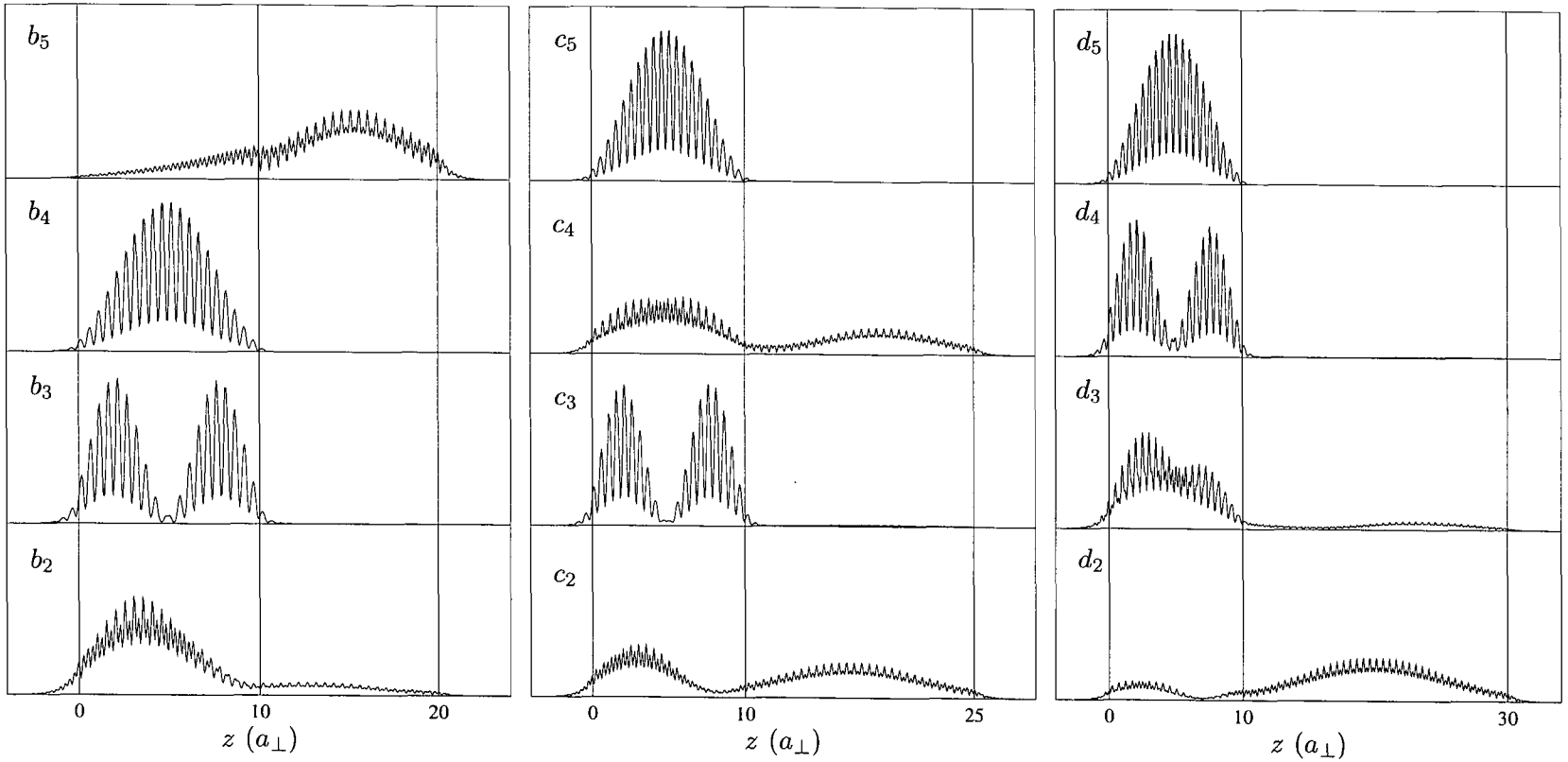
At an InAs width of $30a_{\perp}/2$ (Figure 4.13(b)) we observe hybridization between the light hole and electron ground states, labelled c_2 and c_4, respectively. We also see from Figure 4.11 that this corresponds to the anti-crossing of the light hole and electron energy bands, which is analogous to the minigaps of in-plane band structures. With the InAs layer as wide as $40a_{\perp}/2$ (Figure 4.13(c)) the hybridization of the two states is considerably reduced and we see that the mostly electron-like state, d_2, is now lower in energy than the light hole-like state, d_3, again indicating

that anti-crossing has taken place.

The difference between this present case and those previously reported by other workers is that there is no hybridization with the lowest energy hole state, i.e., the heavy hole ground state. As mentioned earlier, when the InAs layer is sufficiently wide, the electron bands *cross* the ground state heavy hole band. However, minigaps appear in the in-plane band structure with hybridization occurring between electron and heavy hole, as well as light hole states. This can be explained in terms of band mixing within the valence band states for non-zero in-plane wave vector, such that the heavy hole ground state has an appreciable light hole admixture [11] away from the (2D) Brillouin zone centre. At the Brillouin zone centre ($\mathbf{k}_{||} = (0,0)$) the coupling, and hence hybridization, between electron and heavy hole states is negligible[3], hence the lack of anti-crossing and hybridization for these states in Figures 4.11, 4.12 and 4.13.

Energy band anti-crossing as a function of electron wave vector is a familiar phenomenon from bulk energy band structure, and it is interesting to see related effects as the result of a variation in a spatial property of the heterostructure, namely the InAs layer width, rather than the electron wave vector.

[3]In commonly used $\mathbf{k} \cdot \mathbf{p}$ schemes, there is zero coupling between electron and heavy hole states at the zone centre.



(a) InAs width = $20a_\perp/2$ (b) InAs width = $30a_\perp/2$ (c) InAs width = $40a_\perp/2$

Figure 4.13 The variation with InAs layer width of the probability densities of four bound states of the InAs/GaSb/AlSb stepped quantum well, with zero in-plane wave vector. The left hand central region, $0 \leq z \leq 10a_\perp$, is the GaSb layer, with the variable right hand central region consisting of InAs (as in the band line-up diagram, Figure 4.10). The probability densities are plotted according to their corresponding energies, with the highest energy state at the top and the lowest energy state at the bottom of each plot. The labelling is consistent with that of Figure 4.11.

4.3.2 Applied electric field

The application of an electric field to an asymmetric quantum well in the growth direction has fundamentally different effects to its application to a symmetric[4] well. Due to the asymmetry of the unperturbed potential, the charge is already polarised with respect to the centre of the InAs/GaSb region, hence, the applied field does not have to do any work to induce a dipole moment. This results in a linear change in the bound state energies, even for small applied fields. Alternatively, we note that in perturbation theory, the first order linear term in the correction to the unperturbed energy no longer vanishes for small fields since the unperturbed wavefunction is not symmetric. The lack of symmetry in the unperturbed potential also means that the change in energy is no longer symmetric with the applied field but is a monotonic function of the field.

While the characteristics described above are reasonably well known for asymmetric wells in general, the broken gap in this system can result in interesting variations on the underlying linear energy response. Figure 4.14 plots the variation in the 'averaged' energy of the bound states for a stepped well of 20 GaSb and 20 InAs monolayers. By 'averaged' energy, we mean, as in Section 4.2.5, that the electrostatic potential energy origin is set to be midway between the barriers and the bound state energy reference is equivalent to the top of the GaSb valence band at this point. For this particular asymmetric well, the averaged energies must be treated with more caution than for the single quantum well. This is because the hole and electron bound states are concentrated in different spatial regions. The electron states are concentrated in the InAs region, the holes in the GaSb region, and thus the energy separation of electron and hole states *increases* with increasing applied field, the energy separation increasing by approximately $|eV_a/2|$.

[4]Or 'nearly' symmetric, as in the case of the InAs/AlSb single quantum well described in Section 4.2.

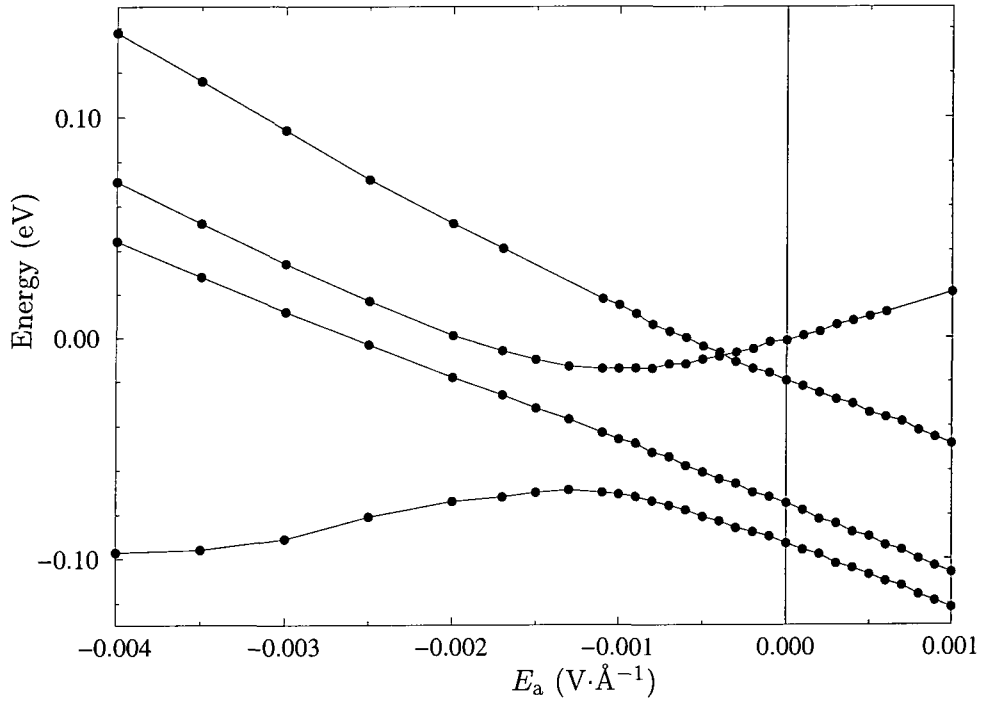


Figure 4.14 The 'averaged' variation with applied electric field of the bound state energies of the asymmetric stepped well heterostructure. Both GaSb and InAs regions are 20 monolayers wide.

Comparison with Figure 4.11 reveals that for zero applied field, the energy levels plotted are, in order of *increasing* energy, lh_1, hh_2, hh_1 and c_1. For a positive E_a, we obtain the expected results, i.e., the energy levels vary linearly, with the electron-hole energy separations increasing. However, for negative E_a the situation is more complex. The energy response remains linear in E_a and eventually the ground state electron and heavy hole levels cross. With an increasingly negative field, the electron and ground light hole states anti-cross in the same energy region in which the electron and first heavy hole excited state cross. The now electron-like level continues to drop but then starts to anti-cross with an excited hole state below the plotted energy range. These effects are similar to those seen in Section 4.3.1, where the crossing and anti-crossing behaviour was a result of variations in the InAs layer width, rather than variations in an applied electric field. So, again we see behaviour analogous to that which causes the minigaps in the in-plane band

structure. Wong *et al.* [91] have reported the results of microscopic calculations on a similar $InAs/Al_{0.1}Ga_{0.9}Sb$ system, including the effects of an applied electric field. However, they do not report any anti-crossing behaviour of the type seen in Figure 4.14. Assuming the calculations were performed using step sizes between different applied fields that are small enough to properly resolve anti-crossing behaviour, its absence is probably the result of spatial separation of the electron and hole states, due to the nature of the multiple quantum well structure studied. They cite this spatial separation to account for the calculated optical matrix elements, which sometimes are much smaller than expected due to little overlap of the wavefunctions involved.

It should be noted that a more rigorous analysis of the states around the band overlap should employ at least some element of self-consistency in the calculations. If accurate predictions were to be made concerning the functioning of devices based on this type of system, band bending effects due to the possibly large scale interband charge transfer between InAs and GaSb layers would have to be taken into account.

4.4 Summary

This chapter has demonstrated the power of the complex band structure EPM technique, described in Chapter 3, for calculating various electronic properties of semiconductor heterostructure bound states. In particular, heterostructures with unusual band profiles cause no problem for this approach.

In Section 4.2, we compared the predicted bound state energies to both calculated and experimentally determined values available in the literature and found satisfactory agreement, suggesting that the subsequent analysis is reasonably reliable. The existence of transitions within the conduction band well at an energy corresponding to the wave length used in fibre optic telecommunications was predicted. Since

intraband transitions between ground and first excited states have a larger TM mode momentum (or, equivalently, dipole) matrix element, and a substantially increased joint density of states than TM or TE mode interband transitions, devices based around this structure may have advantages over existing technology. Also, the variation in the transition energy was shown to be small over a range of the applied electrostatic fields necessary for devices to function.

Section 4.3 contains a study of the InAs/GaSb/AlSb asymmetric well. Again, the complex band structure EPM approach proved very powerful in the analysis of this system. We demonstrated band crossing and anti-crossing behaviour as a function of both the heterostructure dimensions and an applied electrostatic field. The anti-crossing behaviour is claimed to be analogous to the appearance of minigaps in the in-plane band structure, although in this case, the electron states *cross* the heavy hole ground state, since for zero in-plane wave vector the coupling between conduction and heavy hole bands is negligible.

Chapter 5

Envelope function theory and methods

5.1 Introduction

5.1.1 Envelope functions

Envelope function theories involve expressing the electron wavefunction (or pseudopotential wavefunction, in the present work) explicitly in terms of functions that vary on the atomic scale and have the periodicity of the crystal lattice, and those that vary slowly on the atomic scale. A well known example is the Bloch form for the energy eigenfunction of an electron in a bulk crystal (see, for example, p133 of [52])

$$\psi_{m\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{m\mathbf{k}}(\mathbf{r}), \tag{5.1}$$

where m is the band index, $\mathbf{k}$ the electron wave vector and $u_{m\mathbf{k}}(\mathbf{r})$ is a Bloch periodic function, which is periodic with the lattice[1]. In the reduced zone scheme, the wave

[1]Note the distinction between the Bloch function, $\psi_{m\mathbf{k}}$, and the Bloch periodic function, $u_{m\mathbf{k}}$.

vector $\mathbf{k}$ in (5.1) is restricted to the first Brillouin zone (1BZ) and the exponential term changes slowly over a unit cell — the atomic scale. Hence, the exponential term modulates the much more rapidly varying $u_{m\mathbf{k}}(\mathbf{r})$, and imposes an overall 'envelope'.

The wavefunction of a general electron state, such as that in a heterostructure, with energy E, will be some linear combination of these Bloch functions [54],

$$\Psi(\mathbf{r}) = \sum_{m\mathbf{k}} a_{m\mathbf{k}} \psi_{m\mathbf{k}}(\mathbf{r}), \tag{5.2}$$

where the summations are over those values of m and $\mathbf{k}$ corresponding to energy E. The $u_{m\mathbf{k}}$, can be expanded using an orthonormal basis of the Bloch periodic functions chosen to have a particular wave vector, $\mathbf{k}_0$, due to the common periodicity of the two sets of functions,

$$u_{m\mathbf{k}}(\mathbf{r}) = \sum_{n} b_{n\mathbf{k}_0,m\mathbf{k}} u_{n\mathbf{k}_0}(\mathbf{r}). \tag{5.3}$$

Hence, we may rewrite (5.2) as

$$\begin{aligned}
\Psi(\mathbf{r}) &= \sum_{m\mathbf{k}} a_{m\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{n} b_{n\mathbf{k}_0,m\mathbf{k}} u_{n\mathbf{k}_0}(\mathbf{r}) \\
&= \sum_{n} \left(\sum_{\mathbf{k}} \left[\sum_{m} a_{m\mathbf{k}} b_{n\mathbf{k}_0,m\mathbf{k}} \right] e^{i\mathbf{k}\cdot\mathbf{r}} \right) u_{n\mathbf{k}_0}(\mathbf{r}) \\
&= \sum_{n} \left(\sum_{\mathbf{k}} c_{n\mathbf{k}_0,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \right) u_{n\mathbf{k}_0}(\mathbf{r}) \\
&= \sum_{n} F_{n\mathbf{k}_0}(\mathbf{r}) u_{n\mathbf{k}_0}(\mathbf{r}),
\end{aligned} \tag{5.4}$$

where $c_{n\mathbf{k}_0,\mathbf{k}} = \sum_m a_{m\mathbf{k}} b_{n\mathbf{k}_0,m\mathbf{k}}$ and $F_{n\mathbf{k}_0}(\mathbf{r}) = \sum_{\mathbf{k}} c_{n\mathbf{k}_0,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}$ is the most general form of the envelope function.

Of course, in bulk materials the charge density must be periodic with the lattice, since each unit cell is identical. Hence the exponential in (5.1) is merely a phase

factor, i.e., $\mathbf{k} \in \mathbb{R}^3$. However, in the presence of an interface, and therefore in all heterostructures, the wave vector is in general a complex quantity [54,60] and the envelope part of the wavefunction is no longer just a phase factor, but modulates the charge density.[2]

5.1.2 Benefits and problems of envelope function methods

The usefulness of thinking in terms of envelope and periodic functions becomes apparent when introducing approximations to facilitate the solution of the Schrödinger equation, especially in the study of heterostructures. Indeed, probably the most widely used techniques for modelling heterostructures (for example [18,92,93]) are based on Kane's $\mathbf{k} \cdot \mathbf{p}$ model [94,95] for calculating the energy dispersion in bulk crystals, using the symmetry properties of the zone centre periodic functions. Even simpler 'textbook' approaches are capable of producing reliable results in certain circumstances, as demonstrated by the comparison between the EPM and 1D single parabolic band (SPB) results in Chapter 4. These techniques are grouped together under the term 'envelope function approximations' (EFAs), which share the common feature of not requiring explicit knowledge of the Bloch periodic functions, in contrast to pseudopotential and tight-binding methods, where the Bloch periodic functions are a natural product of the calculations. Instead, the information that characterises a material is usually represented by an 'effective mass' for each band included and the energy separation of the bands, both at $\mathbf{k}_0$. Thus, such methods are also referred to as effective mass approximations. We note that EFAs replace the true Schrödinger equation, the solution of which necessarily involves Bloch periodic components, with a 'Schrödinger-like' equation that (usually) has a piece-wise constant potential term, and the energy eigenstates of which are just envelope functions

[2] Although $\mathbf{k}$ is now complex, the real part is still chosen to lie within the 1BZ so that the oscillatory nature of $e^{i\mathbf{k}\cdot\mathbf{r}}$ is still slow on the atomic scale. This may be heavily attenuated if $\mathbf{k}$ has a large imaginary part.

(see, for example, [92]).

EFAs have proven very successful in calculating many properties of heterostructures [22,70] and also in device modelling, usually being implemented in ways requiring little computational power when compared to other methods. However, until recently there has been genuine concern over the physical and mathematical justifications for the use of these EFAs. In particular, the choice of the appropriate boundary conditions to apply to envelope functions at heterojunctions was a contentious issue, with a particular choice being justified by the apparent success of the method in predicting some physical properties rather than by the use of first principles theory. These issues have now been theoretically resolved by Burt [17,96,97], with subsequent contributions from other workers [15,98–100], most notably Foreman [14,16,101].

Burt showed how the wavefunction of an electron in a heterostructure may be expressed in terms of a unique set of envelope functions, which are both continuous *and* smooth over all space, including at abrupt heterojunctions. This is achieved, in part, by defining a single expression for the envelope function expansion of the wavefunction, valid for both well and barrier regions, with the same set of Bloch periodic functions being used throughout, by definition. Furthermore, he derived the usual effective mass equations from the exact Schrödinger equation, analysing the approximations made, and thus formally justifying such methods and resolving the boundary condition quandary. A full review of this work can be found in [17].

5.1.3 Chapter aims

Although Burt's theory is rigorous and developed for application to semiconductor heterostructures, all his numerical results are for idealised one-dimensional models, intended as illustrations of the basic physical principles. Gómez *et al.* [100] have used a 1D implementation of Burt's theory to make comparisons with standard effective

mass methods. They looked at more varied structures than in Burt's original work, including double and graded-potential wells, but these still represented model, rather than realistic, systems.

This chapter describes a method that the author has developed to extract envelope functions from pseudopotential complex band structure calculations with the intention of verifying Burt's theory when applied to realistic three-dimensional (3D) quantum well systems (with the inclusion of spin orbit coupling effects). To the author's knowledge, there have been no other attempts to verify Burt's theory by directly extracting rigorously defined envelope functions from a microscopic model, such as the EPM. Cuypers and van Haeringen [102,103] have defined envelope functions in terms of pseudopotential wavefunctions, but in such a way that sometimes there is more than one envelope function associated with a given band. In addition, their envelope functions are derived from the bulk properties of each heterostructure material, so that the basis states (the Bloch periodic functions) vary from one material to another, in contrast to Burt's theory.

The subsequent chapter examines the results in detail, especially the spatial gradient of the envelope function near an abrupt interface, and investigates some of the consequences of approximations commonly made when using effective mass envelope functions.

5.2 Theory

The starting point for envelope function theory is the fact that the electron wavefunction can be expanded using a complete set of orthonormal functions, $\{u_n(\mathbf{r})\}$, that are periodic with the crystal lattice, as shown in (5.4),

$$\Psi(\mathbf{r}) = \sum_n F_n(\mathbf{r}) u_n(\mathbf{r}), \tag{5.5}$$

where the $F_n(\mathbf{r})$ are defined as the envelope functions, which clearly depend on the particular choice of $\{u_n(\mathbf{r})\}$, which is common throughout the heterostructure.[3] The reference to the wave vector, $\mathbf{k}_0$, of the periodic states has been dropped for notational convenience. In the present work, bulk zone centre states will be used for the periodic functions, so n refers directly to the band index. This is the most common basis for $\mathbf{k} \cdot \mathbf{p}$ calculations as one is usually interested in states 'close' to the zone centre. Only lattice matched systems are considered, so that the $u_n(\mathbf{r})$ straightforwardly form a basis for any function periodic with the lattice in both well and barrier materials. In general, both Ψ and u_n will be spinors, and the calculations presented in this thesis *were* performed with spin included, though the theory presented does not include spinors for notational clarity.

Burt has shown [97] that in order to make the envelope functions unique, it is necessary, when expanding them in plane waves, that only wave vectors in a single Brillouin zone are used — for this work the first Brillouin zone (1BZ) is used.

5.2.1 Supercell approach

Following Burt's method, we require an expression for the wavefunction, in terms of a plane wave expansion, that is valid throughout the quantum well. The method so far expounded in Chapters 3 and 4 to calculate quantum well wavefunctions results in expressions that are valid only within each material layer, since they include summations over the bulk states of each material. Furthermore, that analysis employed the complex band structure approach, and clearly complex wave vectors do not lie solely in the (real) 1BZ. A non-zero imaginary part to the wave vector corresponds to a wavefunction with exponential growth or decay, which may be expressed as a Fourier expansion with an infinite range of real wave vectors.

[3]Here, we consider only lattice matched systems, in which case the lattice parameter is constant along the heterostructure growth direction, though it is possible to include strained materials (see p99 of [104]).

To achieve the required plane wave expansion, a supercell method [97] is employed. The supercell is defined as follows:

- In the z-direction (i.e., the growth direction), the supercell has the quantum well at its centre, and extends far enough into both barrier regions that the electron probability density of a bound state has decayed to a negligible value at the supercell boundaries.
- In the xy-plane, the supercell has dimensions $a \times a$, where a is the lattice constant.

Thus, periodic boundary conditions can be applied along the z-direction, in analogy to the Born-von Karman boundary conditions used for bulk materials. In the plane of the well the wavefunction is already in the required form, since we consider states with real in-plane wave vector components, $\mathbf{k}_{||}$, which should be chosen to lie in the two-dimensional first Brillouin zone ($1\mathrm{BZ}_{||}$). Thus, the wavefunction for the quantum well heterostructure, for a given $\mathbf{k}_{||}$, can be expressed as a Fourier series

$$\Psi_{\mathbf{k}_{||}}(\mathbf{r}) = e^{i\mathbf{k}_{||}\cdot\mathbf{r}_{||}} \sum_k \sum_{\mathbf{g}} a_{\mathbf{g}}(\mathbf{k}_{||}, k) e^{i(\mathbf{e}_z k + \mathbf{g})\cdot\mathbf{r}} \tag{5.6}$$

$$= e^{i\mathbf{k}_{||}\cdot\mathbf{r}_{||}} \phi_{\mathbf{k}_{||}}(\mathbf{r}), \tag{5.7}$$

where $\mathbf{g}$ are reciprocal lattice vectors, $\mathbf{e}_z$ is a unit vector in the z-direction and a 'reduced' wavefunction, $\phi_{\mathbf{k}_{||}}(\mathbf{r})$, has been defined. This is still a 3D function in both direct and reciprocal space but depends on the in-plane wave vector only through the parametric dependence of the Fourier coefficient, $a_{\mathbf{g}}(\mathbf{k}_{||}, k)$. It is the use of this reduced wavefunction that allows the in-plane supercell dimensions to be $a \times a$ for any $\mathbf{k}_{||}$, since $\phi_{\mathbf{k}_{||}}(\mathbf{r})$ is periodic with the lattice in the xy-plane. So, the Fourier

coefficients in (5.6) are given by

$$a_{\mathbf{g}}(\mathbf{k}_{\|}, k) = \frac{1}{\Omega_{sc}} \int_{\Omega_{sc}} \phi_{\mathbf{k}_{\|}}(\mathbf{r}) e^{-i(\mathbf{e}_z k + \mathbf{g})\cdot\mathbf{r}} d^3\mathbf{r}, \tag{5.8}$$

with Ω_{sc} being the volume of the supercell, and the 3D wave vector, $\mathbf{k} = (\mathbf{k}_{\|}, k)$, lying in the 1BZ. The zone centre states can also be expressed as a Fourier series,

$$u_n(\mathbf{r}) = \sum_{\mathbf{g}} u_{n\mathbf{g}} e^{i\mathbf{g}\cdot\mathbf{r}}, \tag{5.9}$$

with the Fourier coefficients given by

$$u_{n\mathbf{g}} = \frac{1}{\Omega_c} \int_{\Omega_c} u_n(\mathbf{r}) e^{-i\mathbf{g}\cdot\mathbf{r}} d^3\mathbf{r}, \tag{5.10}$$

where Ω_c is the unit cell volume. Further, the zone centre functions are chosen to be orthonormal. Thus, the set of plane waves with reciprocal lattice wave vectors, $\{e^{i\mathbf{g}\cdot\mathbf{r}}\}$, can be expanded in terms of zone centre states, allowing the wavefunction to be expressed in the form [17,97]

$$\Psi_{\mathbf{k}_{\|}}(\mathbf{r}) = \sum_{n} \left(\sum_{k} \sum_{\mathbf{g}} a_{\mathbf{g}}(\mathbf{k}_{\|}, k) u^*_{n\mathbf{g}} e^{ikz} e^{i\mathbf{k}_{\|}\cdot\mathbf{r}_{\|}} \right) u_n(\mathbf{r}), \tag{5.11}$$

which, after comparison with (5.5), yields an expression for the envelope functions

$$F_n(\mathbf{r}) = \sum_{k} \sum_{\mathbf{g}} a_{\mathbf{g}}(\mathbf{k}_{\|}, k) u^*_{n\mathbf{g}} e^{ikz} e^{i\mathbf{k}_{\|}\cdot\mathbf{r}_{\|}}. \tag{5.12}$$

Finally, the $\mathbf{r}_{\|}$ dependence can be factored out of (5.12) to define an envelope function that is a function of z only,

$$f_n(z) = \sum_{k} \sum_{\mathbf{g}} a_{\mathbf{g}}(\mathbf{k}_{\|}, k) u^*_{n\mathbf{g}} e^{ikz}. \tag{5.13}$$

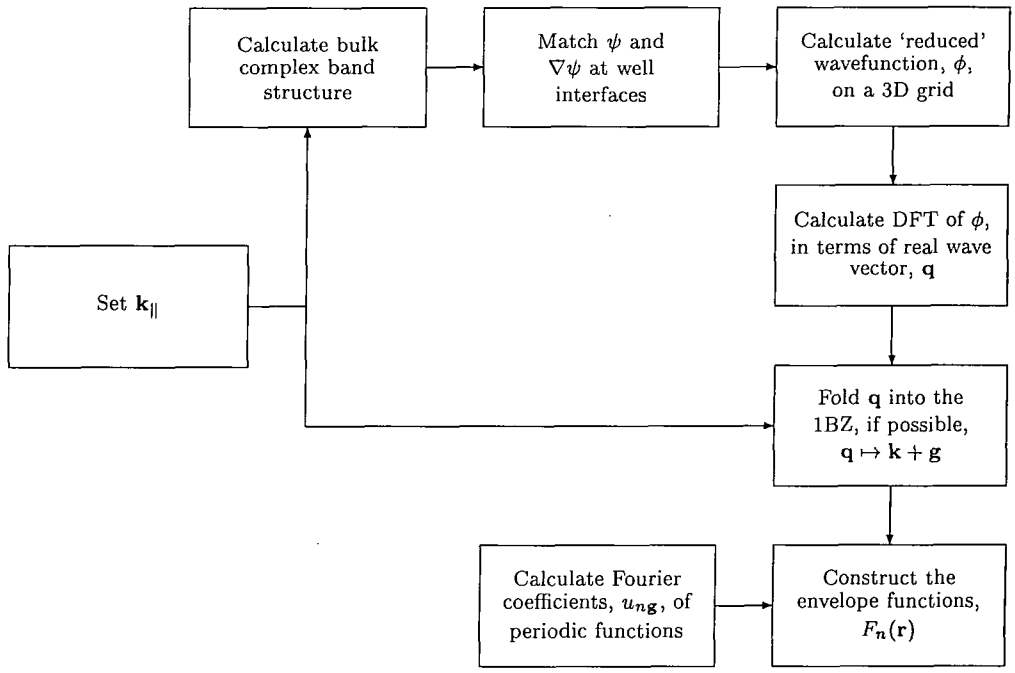


Figure 5.1 The main steps involved in deriving the envelope functions, as described in the text, from wavefunctions generated using the complex band structure matching technique.

5.3 Implementation

The preceding section has dealt with the theoretical derivation of a set of unique and continuous envelope functions for a 3D supercell. This section illustrates how one calculates these functions in practice, starting with the results of the complex-k wavefunction matching technique described in Chapter 3. The main steps are outlined in Figure 5.1.

5.3.1 Wavefunction calculation

If the complex band structure approach has been used to obtain energy eigenfunctions of quantum well bound states, it is not possible to apply (5.12) directly to obtain the envelope functions. This is because the sum over k in (5.12) must only include those z-components for which $\mathbf{k} = (\mathbf{k}_{||} + \mathbf{e}_z k)$ is in the 1BZ, and clearly a complex or purely imaginary component does not satisfy this condition, as men-

tioned in Section 5.2.1.

However, by imposing the supercell periodicity, it is possible to express the wavefunction as in (5.6), with real $\mathbf{k} \in$ 1BZ. The first step is to calculate the value of the wavefunction on a 3D grid throughout the supercell, using the appropriate complex k and expansion coefficients in each region, and then take a discrete Fourier transform (DFT). In fact, it is more convenient to calculate the DFT of the 'reduced' wavefunction, $\phi_{\mathbf{k}_\parallel}(\mathbf{r})$, defined in Section 5.2.1. This yields a set of Fourier coefficients, $a_\mathbf{q}$, of the wavefunction, where $\mathbf{q}$ is the *real* DFT wave vector,

$$a_\mathbf{q} = \frac{1}{N} \sum_\mathbf{r} \psi(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} \tag{5.14}$$

$$\phi_{\mathbf{k}_\parallel}(\mathbf{r}) = \sum_\mathbf{q} a_\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{r}}, \tag{5.15}$$

where N is the total number of samples, and $\mathbf{q}$ and $\mathbf{r}$ are now *discrete* variables taking N values (see Section 5.3.3). The wave vector $\mathbf{q}$ can always be split into a sum of a vector in the 1BZ and a reciprocal lattice vector,

$$\begin{aligned} \mathbf{q} &= \mathbf{k}_\parallel + \mathbf{e}_z k + \mathbf{g}, \\ &= \mathbf{k} + \mathbf{g}. \end{aligned} \tag{5.16}$$

For a given $\mathbf{q}$, there will be a unique $\mathbf{k}$ and $\mathbf{g}$, so, for each $\mathbf{q}$ there is a unique mapping

$$a_\mathbf{q} \mapsto a_\mathbf{g}(\mathbf{k}_\parallel, k). \tag{5.17}$$

Band	Basis states
Conduction	$\|S\uparrow\rangle$, $\|S\downarrow\rangle$
Heavy Hole	$\sqrt{\frac{1}{2}}\|(X+iY)\uparrow\rangle$, $\sqrt{\frac{1}{2}}\|(X-iY)\downarrow\rangle$
Light Hole	$\sqrt{\frac{1}{6}}\left(\|(X+iY)\downarrow\rangle - \sqrt{\frac{2}{3}}\|Z\uparrow\rangle\right)$, $-\sqrt{\frac{1}{6}}\left(\|(X-iY)\uparrow\rangle - \sqrt{\frac{2}{3}}\|Z\downarrow\rangle\right)$
SSO	$\sqrt{\frac{1}{3}}\left(\|(X+iY)\downarrow\rangle + \|Z\uparrow\rangle\right)$, $-\sqrt{\frac{1}{3}}\left(\|(X-iY)\uparrow\rangle - \|Z\downarrow\rangle\right)$

Table 5.1 Typical $\mathbf{k}\cdot\mathbf{p}$ basis states, expressed in terms of S and P-orbitals. These states are simultaneous eigenvectors of the Hamiltonian and the total angular momentum operator, $\widehat{H}$ and $\widehat{J}$, respectively.

5.3.2 Zone centre states

The zone centre coefficients, $u_{n\mathbf{g}}$, have been found in the usual way by calculating the eigenvectors of the (pseudo) Hamiltonian at the Γ-point of the well material. It is important for the derivation of the envelope function expression in (5.12) to ensure that these states are orthogonal. It is also useful, for the purposes of a direct comparison of individual envelope functions with $\mathbf{k}\cdot\mathbf{p}$ results, to choose zone centre states that are also eigenvectors of the total angular momentum.[4] The majority of $\mathbf{k}\cdot\mathbf{p}$ calculations include at most the spin split off (SSO), light, heavy and first conduction bands and the relevant symmetries of the basis states used are listed in Table 5.1.

5.3.3 Application of discrete Fourier transforms

The discrete Fourier transform (DFT) of a 3D function, φ, periodic over a finite volume (in this case the supercell) with values φ_{s_1,s_2,s_3}, is defined as[5] (see for example

[4]By convention, the angular momentum is quantised in the z-direction. This choice is abitrary but it is important to note that the form of the Pauli spin matrices used in the EPM calculations (see Section 2.2.3) determines this direction. One can still have the well growth along either x or y, but the wavefunction cannot then also be made an eigenvector of $\widehat{J}$.

[5]With this definition, no factor of N^{-1} appears in the expression for the inverse DFT. Many authors use $N^{-\frac{1}{2}}$ for both direct and inverse DFTs.

p91 of [105])

$$\tilde{\varphi}_{t_1,t_2,t_3} = \frac{1}{N}\sum_{s_1}\sum_{s_2}\sum_{s_3}\varphi_{s_1,s_2,s_3}e^{-i2\pi\left(\frac{t_1 s_1}{N_1}+\frac{t_2 s_2}{N_2}+\frac{t_3 s_3}{N_3}\right)} \tag{5.18}$$

where the s_j and t_j are integers labelling the discrete position in direct and reciprocal space, respectively, N_j is the number of samples in the j^{th} direction, i.e., the number of positions at which the wavefunction is calculated, and $N = N_1 N_2 N_3$. The summations are over the values

$$s_j = -\left(\frac{N_j}{2}-1\right), -\left(\frac{N_j}{2}-1\right)+1, \ldots, -1, 0, 1, \ldots, \frac{N_j}{2}-1, \frac{N_j}{2}, \tag{5.19}$$

with the possible values for the t_j similarly defined.

Consider a supercell of dimension $M_1 \times M_2 \times M_3$, in lattice constants, a. The position, x_j, along the j^{th} direction is defined by s_j through[6]

$$x_j = \frac{s_j}{N_j}M_j a, \qquad s_j = -\left(\frac{N_j}{2}-1\right), \ldots, -1, 0, 1, \ldots, \frac{N_j}{2}, \tag{5.20}$$

and hence,

$$\frac{s_j}{N_j} = \frac{x_j}{M_j a}. \tag{5.21}$$

The exponent in (5.18) becomes

$$-i\frac{2\pi}{a}\left(\frac{t_1 x_1}{M_1}+\frac{t_2 x_2}{M_2}+\frac{t_3 x_3}{M_3}\right) = -i\mathbf{q}\cdot\mathbf{r}, \tag{5.22}$$

[6]Since it is assumed that the two materials forming the quantum well are lattice matched, we have $a_j = a$, for $j = 1, 2, 3$.

where we have introduced a 'DFT wave vector', $\mathbf{q}$, given by

$$\mathbf{q} = \frac{2\pi}{a}\left(\frac{t_1}{M_1}, \frac{t_2}{M_2}, \frac{t_3}{M_3}\right). \tag{5.23}$$

So, if $\mathbf{q}$ and $\mathbf{r}$ are expressed in units of $2\pi/a$ and a, respectively, the exponent can be written

$$-i2\pi\left(\frac{t_1 x_1}{M_1} + \frac{t_2 x_2}{M_2} + \frac{t_3 x_3}{M_3}\right) \tag{5.24}$$

where t_j can take the values given in (5.19).

5.3.4 Wave vector folding

Equation (5.23) shows that the j^{th} component of $\mathbf{q}$ has values over the range

$$q_j = \left(-\frac{N_j}{2}+1\right)\frac{1}{M_j}, \left(-\frac{N_j}{2}+2\right)\frac{1}{M_j}, \ldots, -\frac{1}{M_j}, \tag{5.25}$$

$$0, \frac{1}{M_j}, \ldots, \left(\frac{N_j}{2}-1\right)\frac{1}{M_j}, \frac{N_j}{2M_j}\left(\times\frac{2\pi}{a}\right). \tag{5.26}$$

Now, M_j and N_j are the number of lattice constants, a, and data samples in the j^{th} direction, respectively. Clearly, for a good approximation to the *continuous* wavefunction, one should use as many samples per layer as possible. Therefore, we would like to have

$$N_j \gg M_j, \tag{5.27}$$

which means that most of the wave vectors in (5.26) will have components larger than $2\pi/a$, and will therefore lie outside the first Brillouin zone[7]. However, the wave

[7]Of course, some vectors with components less than or equal to $2\pi/a$ will still lie outside the 1BZ, such as $(111)2\pi/a$, if we define the 1BZ as the standard truncated octahedron. However, the largest *possible* component is $1 \times 2\pi/a$.

vector can be split into the sum of a wave vector inside the 1BZ, $\mathbf{k}$, and a reciprocal lattice vector, $\mathbf{g}$, as described in Section 5.3.1. We call this 'wave vector folding', and to avoid verbosity in the following discussion, we say that a Fourier coefficient has been folded, when in fact it is the associated wave vector that has been folded.

Effects of a finite $\{\mathbf{g}\}$

Within the approximation of discretising the wavefunction, the DFT theory is exact. However, in any numerical calculation, the number of reciprocal lattice vectors used must be finite and this has consequences for the wave vector folding procedure. The largest value of $\mathbf{q}$ in the j^{th} direction is $N_j/(2M_j)$ in $2\pi/a$, so, depending on the ratio N_j/M_j there may be $\mathbf{q}$'s that cannot be folded because there will be no $\mathbf{g}$ of large enough magnitude to facilitate the wave vector folding.

In order to avoid wasting computational resources, care must be taken when choosing values of N_j and M_j to satisfy (5.27) when the set of $\mathbf{g}$ is finite. If the ratio N_j/M_j is too large then many of the DFT coefficients will have to be excluded from the envelope function calculation.

In practice, this does not present a serious problem (see Sec. 5.4) and the ratio $N_j/(2M_j)$ is set by the condition

$$\frac{N_j}{2M_j} = g_{\max} + 1, \tag{5.28}$$

where $g_{\max}$ is the largest component of the reciprocal lattice vectors in units of $2\pi/a$. To see why this is chosen, consider the set of 89 $\mathbf{g}$-vectors used in this work. This includes vectors, with increasing magnitude, up to $\{331\}$. However, the largest components come from the $\{400\}$ types and so from (5.28) we have a ratio

of $N_j/(2M_j) = 5$. The reason for the addition of 1 is now clear, since, for example,

$$\mathbf{q} = (5,0,0) = (1,0,0) + (4,0,0) \tag{5.29}$$

$$= \mathbf{k} + \mathbf{g}, \tag{5.30}$$

where $\mathbf{k}$ is in the 1BZ and $\mathbf{g}$ is included in the finite set $\{\mathbf{g}\}$. It should be noted that the size of $\{\mathbf{g}\}$ used in extracting the envelope functions need not be the same as that used in the original complex band structure calculations, as a much larger $\{\mathbf{g}\}$ could be used allowing a greater $N_j/(2M_j)$ ratio. This work uses 89 $\mathbf{g}$'s throughout as a matter of convenience, both in ensuring compatibility between the various programmes needed and in computational effort.

5.4 Error analysis

Before examining the nature of the individual envelope functions it is important to check that the expansion (5.5), with the calculated F_n, does give a good representation of the wavefunction, especially considering that some information is lost in the use of a finite set of reciprocal lattice vectors in the folding procedure.[8] Therefore, in order to demonstrate the accuracy of the envelope function expansion, both the wavefunction and the in-plane averaged probability density have been calculated using a number of levels of approximation.

To examine the effects of using a finite $\{\mathbf{g}\}$ in (5.8) and a supercell of finite z dimension it is sufficient to compare the wavefunction as calculated in (5.6), using only those DFT Fourier coefficients allowed after the folding procedure and taking into account the allowed values of k (which are determined by the number of z-layers in

[8]Remembering that even a very small imaginary k component results in a wavefunction with (real) exponential dependence on position and has a Fourier representation with an infinite number of terms.

the supercell), with that calculated using the original complex band structure. The envelope expansion itself, (5.5), gives identical results to (5.6), since there is no further loss in information.[9] However, by calculating the envelope functions explicitly, it is possible to examine the effect of using only some terms in the expansion (5.5). These comparisons are shown for the electron ground state in an $In_{0.53}Ga_{0.47}As/InP$ system, with a well width of $55a/4$ and III–V$\cdots$V–III interfaces — other systems, including InSb/CdTe and InAs/AlSb, have also been studied with similar results. The $In_{0.53}Ga_{0.47}As/InP$ system forms a standard type I well — the offset used is given in Appendix D. Note, however, that the well is an integer number of *half*-monolayers wide. This is done to produce a microscopically symmetric (in-plane averaged) potential with respect to the centre of the well, and hence bound states with symmetric probability densities. This allows for a more direct comparison with standard envelope function approaches, which model single wells as perfectly symmetric (cf. the discussion in Section 4.2.5).

Figure 5.2 shows the 'true' complex band structure wavefunction, as described in Chapters 3 and 4, along with the absolute error in using (5.6). The data is plotted at the arbitrary in-plane coordinate $\mathbf{r}_{||} = (0.1, 0.2)a$ to ensure that the 3D nature of the original wavefunction is maintained by the envelope function expansion. Clearly, the absolute error is negligible, hence the effects of imposing the supercell boundary conditions, sampling the (continuous) wavefunction and omitting unfoldable wave vectors is minimal. To examine the errors further, they are also plotted with a magnification factor of 1000. This indicates that by far the largest errors are near the interfaces between the two materials. This is to be expected for the following reasons:

- In the well region, states with purely real wave vector (in the complex band

[9]There will of course be small differences due to round off errors, but even for finite $\{\mathbf{g}\}$ and M_3, (5.5) and (5.6) are mathematically equivalent.

structure description) tend to dominate the total wavefunction. Since the same $\{\mathbf{g}\}$ has been used in both the complex band structure and envelope function calculations, these states can be represented without information loss (apart from that due to the discretisation process).

- In the barriers, where there are no real wave vectors, the dominant states have wave vectors with only 'small' imaginary parts. The Fourier representation of these states is therefore sharply peaked about the origin. The use of a finite number of $\mathbf{g}$ when constructing the envelope functions has only a negligible effect (this is also true for exponentially varying componenets of the wavefunction that occur in the well region).

- In the barriers, but away from the interfaces, the wavefunction is small and monotonically decaying, so the absolute errors will be reduced.

In other words, the wave vector folding procedure using a finite $\{\mathbf{g}\}$ causes errors near the interfaces, though these are still extremely small relative to the magnitude of the wavefunction. The wavefunction is plotted again in Figures 5.4 and 5.5. This time the errors are those due to using only a limited number of terms in the envelope function expansion. It is apparent from Figure 5.4 that although the complete set of basis functions, $u_n(\mathbf{r})$, numbers 89 ($\times 2$, including spin), an expansion using only 4 ($\times 2$) bands,

$$\Psi \approx \sum_v F_v u_v + F_c u_c, \qquad (5.31)$$

$$v = \text{sso, lh, hh}$$

is quite adequate in representing the wavefunction. This is because the bound states in the quantum well are much closer to the four band edges than to any other bands.

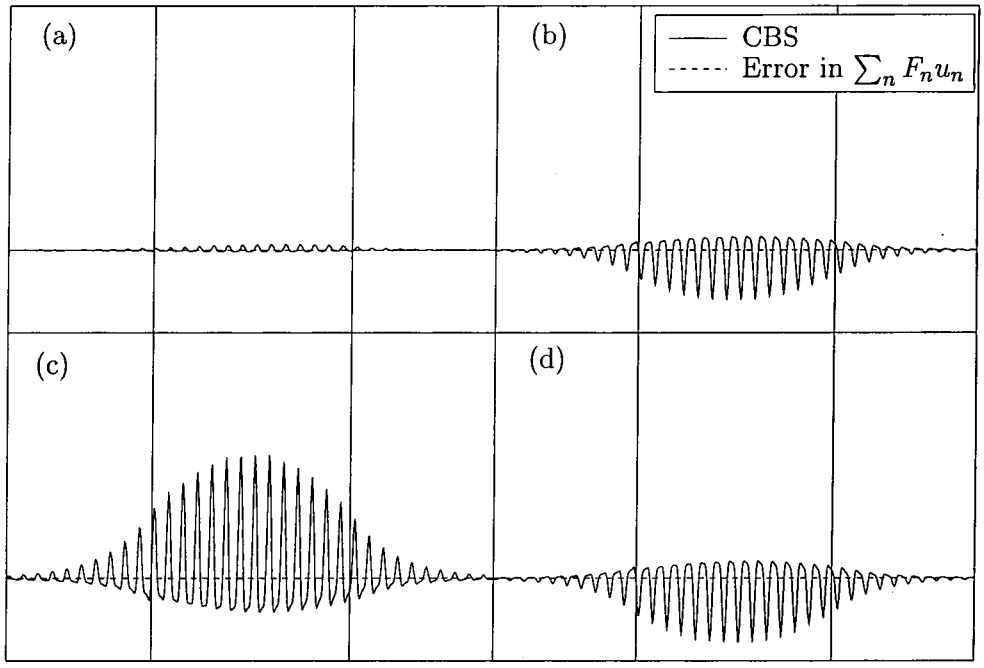


Figure 5.2 'True' CBS wavefunction at $\mathbf{r}_{||} = (0.1, 0.2)a$ and errors in (5.5), or (5.6). (a) Real part, spin up component. (b) Imaginary part, spin up. (c) Real part, spin down. (d) Imaginary part, spin down.

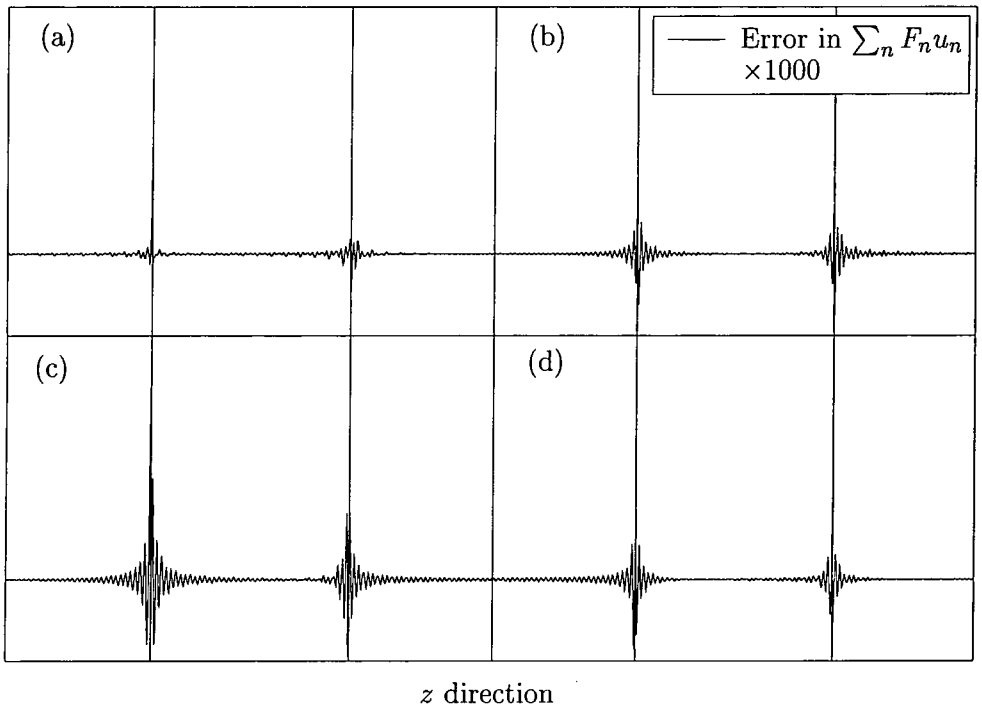


Figure 5.3 Magnified error in the envelope function expansion of the wavefunction plotted in 5.2. (a) Real part, spin up component. (b) Imaginary part, spin up. (c) Real part, spin down. (d) Imaginary part, spin down.

Figure 5.5 shows the errors in using the expansion

$$\Psi \approx F_c u_c, \tag{5.32}$$

which includes only the first (doubly degenerate) conduction band. This crude approximation is still reasonable, with the errors mainly being in the phase of the wavefunction, rather than its shape. However, there are subtle issues involved here, which will be discussed in Chapter 6, and the apparent success of this approximation should be treated with great caution. Results for the analogous approximations for valence band bound states have similarly small errors. The corresponding in-plane averaged probability density is shown in Figures 5.6 and 5.7, the latter using the same approximations as in Figures 5.4 and 5.5. These confirm the previous observations concerning the wavefunctions.

5.5 Example

Having confirmed the validity of the envelope function expansion and the accuracy of the method used in its calculation, we briefly illustrate the results with a typical example, which is discussed in more detail in the following chapter.

Again, the state under consideration is one of the spin degenerate pair of conduction band quantum well ground states of an $In_{0.53}Ga_{0.47}As/InP$, of width $55a/4$. Figure 5.8 shows the z-dependent part of the envelope functions, $f_n(z)$, which dominate in the expansion. These correspond to the bands that are close to the fundamental band gap, as one might expect for a state near the conduction band edge. Since spin is included explicitly in the calculations, each band is doubly degenerate at the zone centre, and hence the zone centre Bloch functions, $u_n(\mathbf{r})$, exist in pairs. For clarity, the envelope function of only one band per pair is shown, and their phases

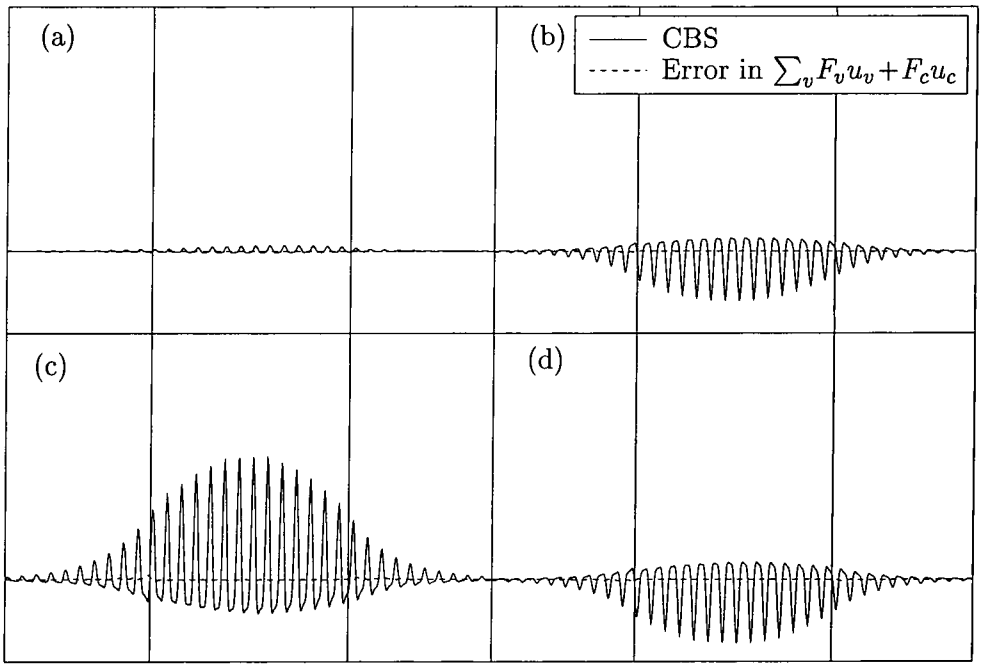


Figure 5.4 'True' CBS wavefunction at $\mathbf{r}_{||} = (0.1, 0.2)a$ and errors in (5.5) using bands 3–10 (including spin). (a) Real part, spin up component. (b) Imaginary part, spin up. (c) Real part, spin down. (d) Imaginary part, spin down.

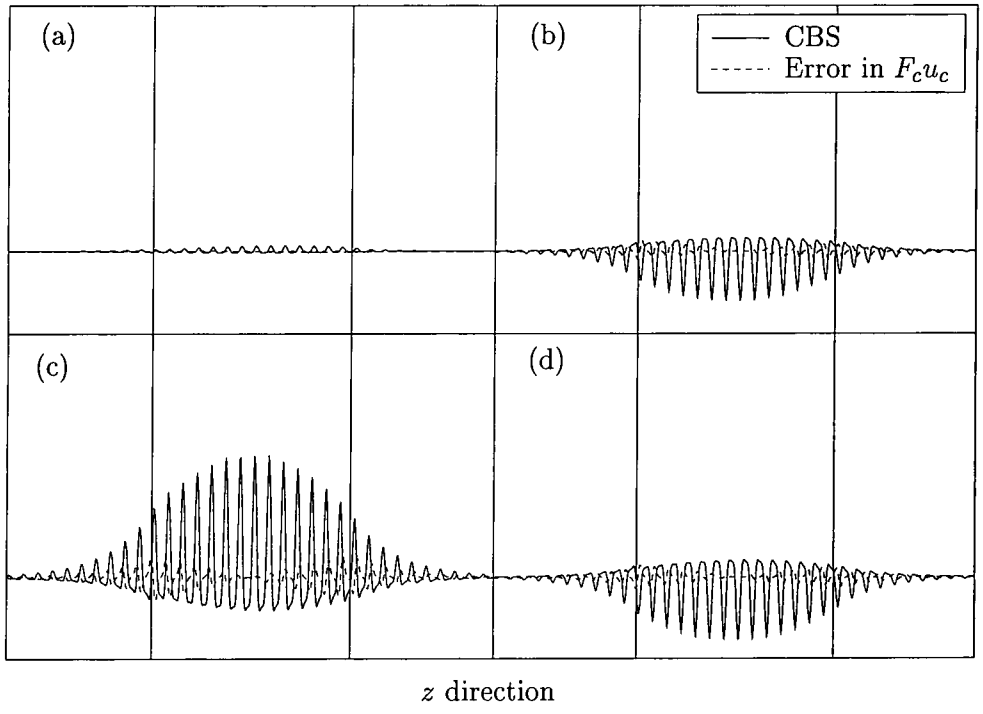


Figure 5.5 'True' CBS wavefunction at $\mathbf{r}_{||} = (0.1, 0.2)a$ and errors in (5.5) using bands 9–10 (including spin). (a) Real part, spin up component. (b) Imaginary part, spin up. (c) Real part, spin down. (d) Imaginary part, spin down.

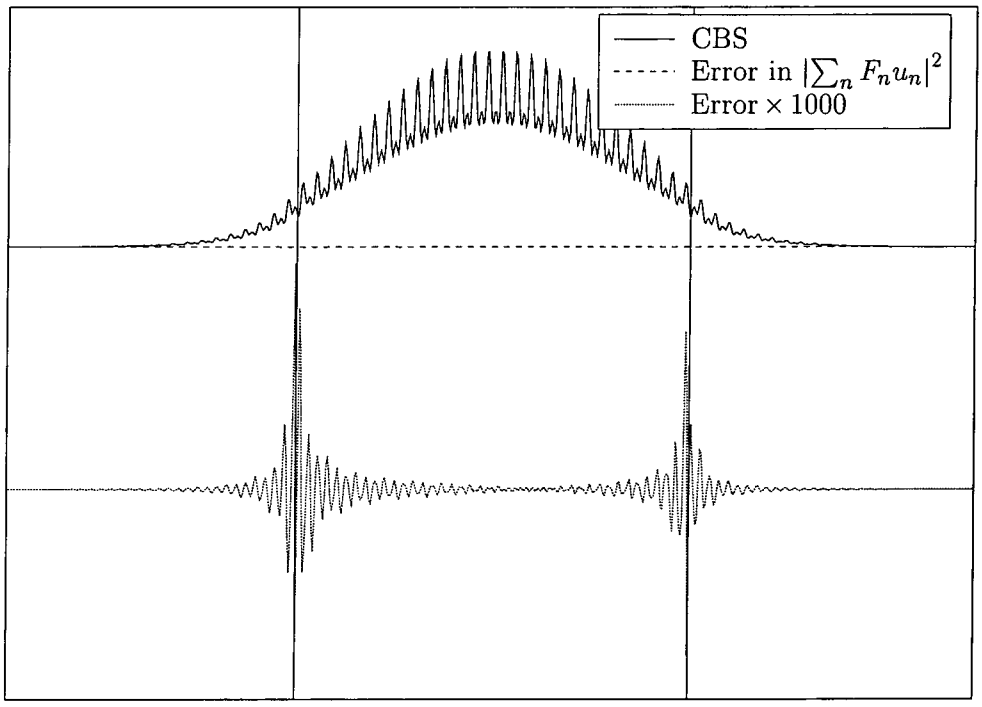


Figure 5.6 In-plane averaged CBS probability density and errors when using (5.5), or (5.6). The magnified error plot has been offset on the vertical axis for clarity.

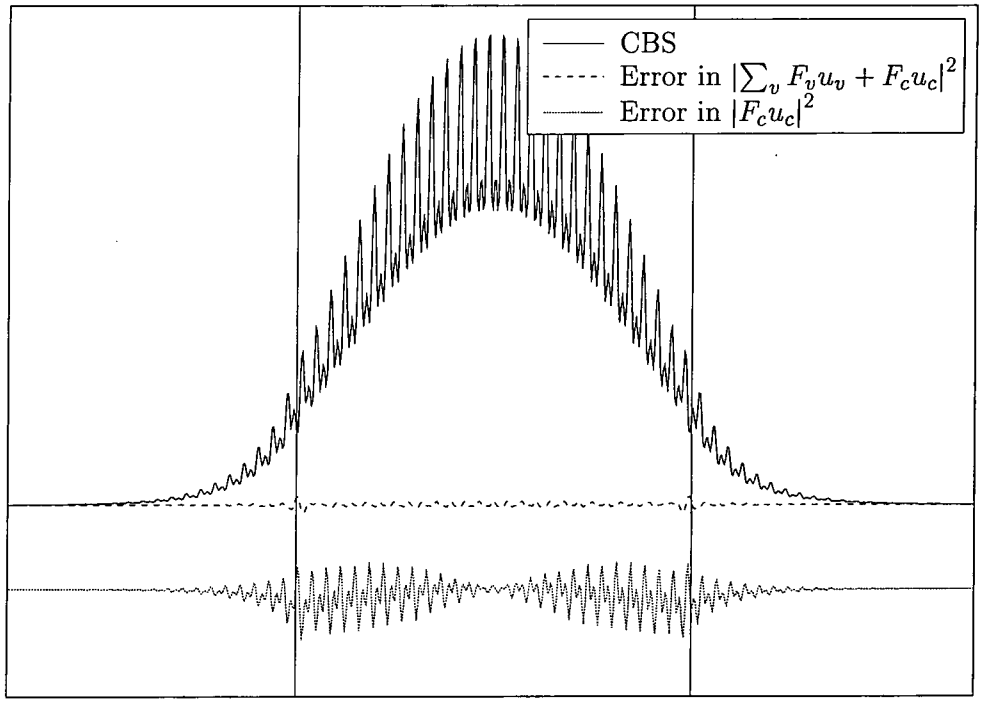


Figure 5.7 In-plane averaged CBS probability density and errors in (5.5) using bands 3–10 and 9–10 (including spin). The error in using $|F_c u_c|^2$ has been offset on the vertical axis for clarity.

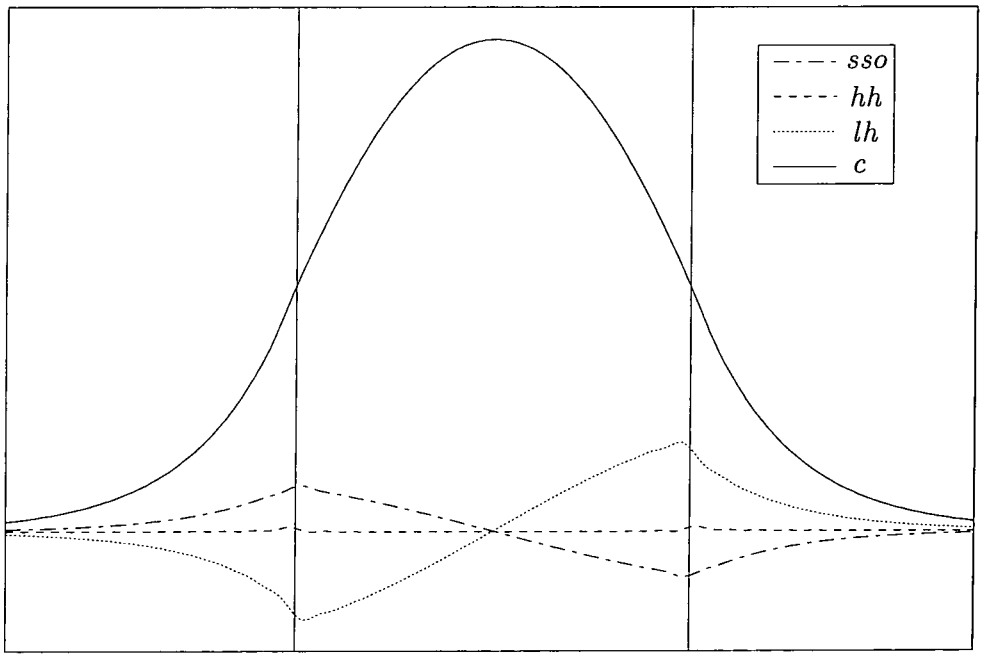


Figure 5.8 Dominant $f_n(z)$ for conduction band single quantum well ground state. See text for details.

have been adjusted to make them real.

5.6 Summary

A technique has been developed that allows the derivation of a unique set of quantum well envelope functions (for a given basis of periodic functions) from realistic three-dimensional pseudopotential complex band structure calculations. Comparisons of the wavefunction expressed in terms of the envelope function expansion with the original data shows that the errors introduced in the process are negligible.

It has also been demonstrated that for states close to the band edges only a few terms dominate the envelope function expansion, (5.5), and that these are due to the band edge periodic functions either side of the fundamental gap (including the spin split-off band). This suggests that EFA schemes that only explicitly include the band edge terms are quite satisfactory. Even the single band edge representation of

the conduction band state,

$$\Psi \approx F_c u_c, \tag{5.33}$$

appears to approximate the wavefunction and charge density reasonably well, justifying its wide use in many situations (but certainly not all situations, as is demonstrated in Sections 6.3.3 and 6.4.2). The error analysis presented in this chapter is for an $In_{0.53}Ga_{0.47}As/InP$ quantum well, but it should be noted that similarly good results hold for the InSb/CdTe system featured in the following chapter.

Chapter 6

Envelope function results and approximations

6.1 Introduction

Chapter 5 explains how quantum well (and other heterostructure) bound state wavefunctions may be expressed in terms of an expansion of well defined envelope functions. Much of the chapter is devoted to a description of a numerical method for calculating the envelope functions from wavefunctions generated using the complex band structure pseudopotential method of Chapter 3. In this chapter, the nature of the envelope functions is investigated in more detail and comparisons made to the results of conventional methods. Of particular interest is the behaviour of the envelope function, and particularly its spatial derivative, in the vicinity of an interface, which until Burt's work (see [17] and references therein), was poorly understood.

The second half of the chapter examines dipole matrix elements between bound states within the quantum well. These are calculated directly from the complex band structure pseudopotential method and using envelope function expansions derived from these. The use of the envelope function expansions allows the results of Burt

[106] on failings of the so called atomic picture of interband dipole matrix elements in a model one-dimensional semiconductor, to be tested with the electronic structure of a real material. Related predictions on charge oscillation induced by coherent interband optical excitation [107,108] are tested in the same way.

6.2 Results

6.2.1 Nature of the envelope functions

Having confirmed in Chapter 5 the validity of the envelope function expansion and the accuracy of the method used in its calculation, it is now appropriate to make a detailed examination of the envelope functions themselves. However, before proceeding, it is convenient to discus some of the complications introduced by the inclusion of spin in the calculations.

- At the zone centre (Γ-point), each energy band is doubly degenerate, i.e., the conduction, heavy hole, light hole, spin split-off and all other bands are each doublets.[1] In this work, the term energy 'band' refers to a doublet (unless stated otherwise). It should therefore be understood that the term 'heavy hole band', for example, is short-hand for the 'pair of degenerate heavy hole bands'.

- Since each band is actually a doublet, there are two degenerate electron states, corresponding to the two spin states of the electron.[2] Thus, there are two zone centre states per band (using the terminology adopted above).

[1] In the absence of strain, there is further degeneracy, such as that of the light and heavy hole bands.

[2] The degenerate electrons must obey the Pauli exclusion principle, however, in general it is not possible for one state to be purely spin up and the other purely spin down.

- So, an envelope function expansion of a quantum well state will include two envelope functions for each band, one for each of the two zone centre states per band.

- The doublet energy degeneracies are not lifted for wave vectors along the primary axes in reciprocal space (the Δ symmetry line between Γ and X).

- Therefore, quantum well states with zero in-plane wave vector will be doubly degenerate, since the growth direction is parallel to a primary axis (the z-axis, in this work).

- For a particular band, the envelope functions of a pair of degenerate bound states, labelled (1) and (2), are related by

$$\left|F_i^{(1)}(\mathbf{r})\right|^2 = \left|F_{i'}^{(2)}(\mathbf{r})\right|^2,$$

  where i and i' label the degenerate zone centre states and we have used notation such that the j^{th} bound state has the envelope function expansion $\sum_i F_i^{(j)}(\mathbf{r})u_i(\mathbf{r})$.

We consider the $55a/4$ $In_{0.53}Ga_{0.47}As/InP$ system introduced in Section 5.4. The z dependent parts, $f_n(z)$, of various envelope functions are plotted in Figures 6.1–6.5, for zero in-plane wave vector. These are for one of the spin degenerate pair of the lowest two electron and lowest three hole bound states, with all the $f_n(z)$ being made real (for clarity) by adjusting their phase. Of the two envelope functions corresponding to each band, only the one with the largest magnitude is shown for clarity.

The key points of interest are outlined below.

- The envelope functions do indeed vary slowly over the unit cell.

- They are continuous *and* smooth — their smoothness is discussed later, with reference to the gradients of the envelope functions near heterojunctions.

- The non-negligible envelopes can be divided into two groups according to which zone centre state they are associated with in the envelope function expansion, i.e.,

  1. Conduction and heavy hole, c and hh, respectively.

  2. Light hole and spin split off, lh and sso, respectively.

  For a given bound state, one group has envelope functions with even parity and the other with odd parity, with respect to the centre of the well. This has been explained within the $\mathbf{k} \cdot \mathbf{p}$ formalism by other workers [12].

- For each bound state, a single band dominates the envelope function expansion. The corresponding envelope has the general form expected from simple 'particle in a box' calculations.

- At zero in-plane wave vector, $\mathbf{k}_{||} = (0, 0)$, the heavy hole band is mostly, but not entirely, decoupled from the other bands. The standard formalism of the $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian has no mixing between the hh band and the other bands at $\mathbf{k}_{||} = (0, 0)$ [109,110].

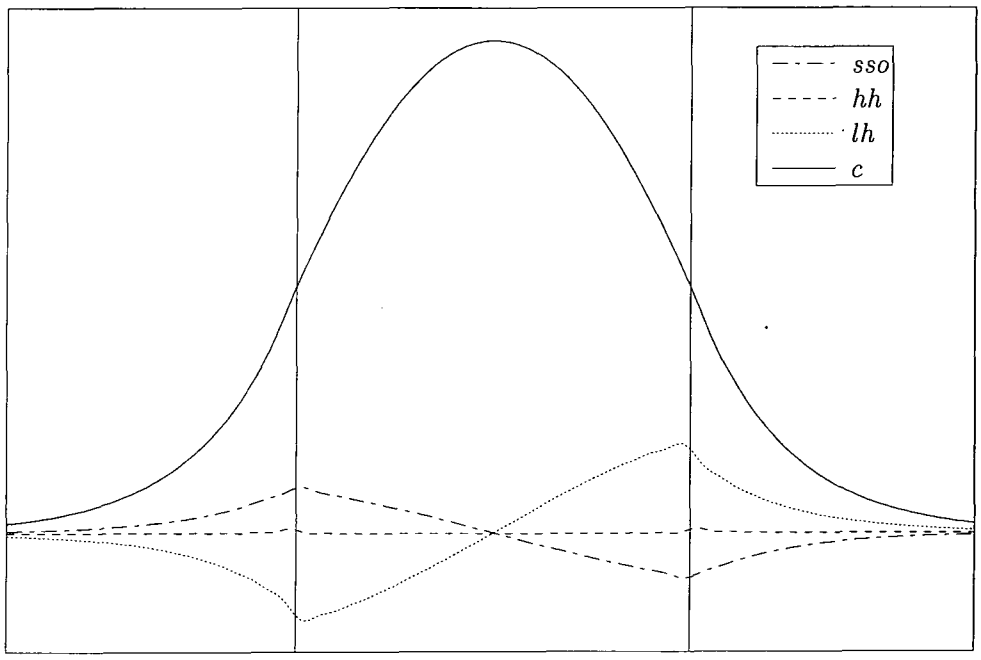


Figure 6.1 Dominant $f_n(z)$ for electron ground state, at $\mathbf{k}_{||} = (0,0)$. The well interfaces are indicated by vertical lines.

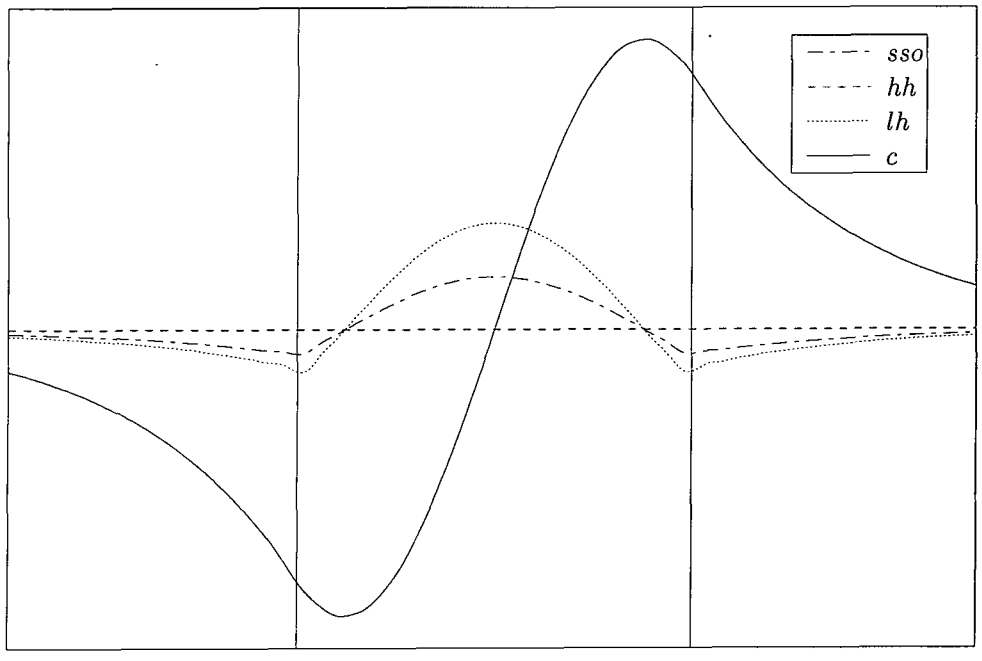
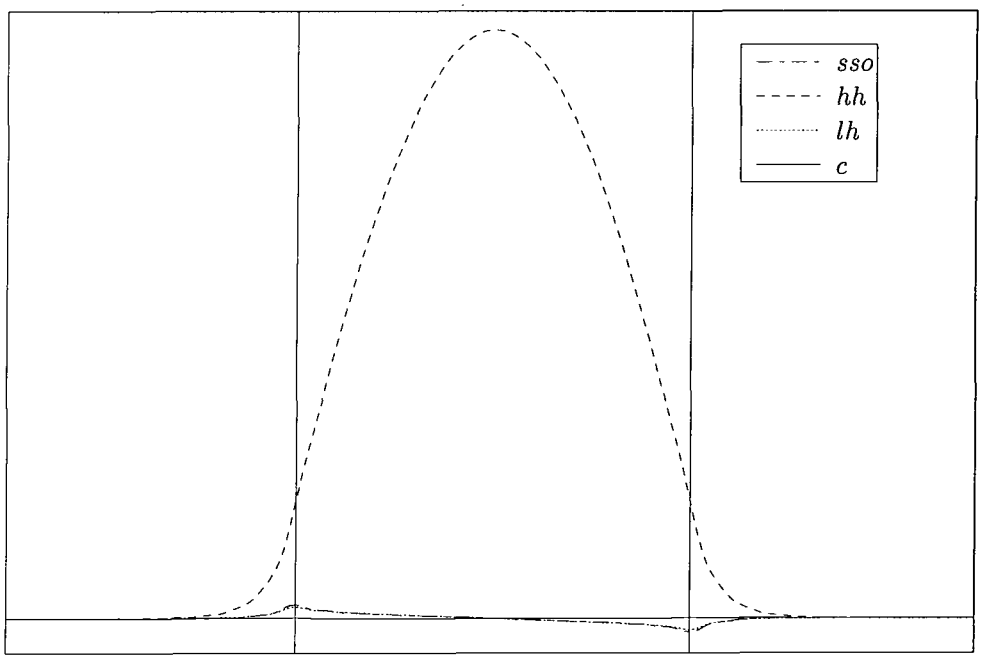
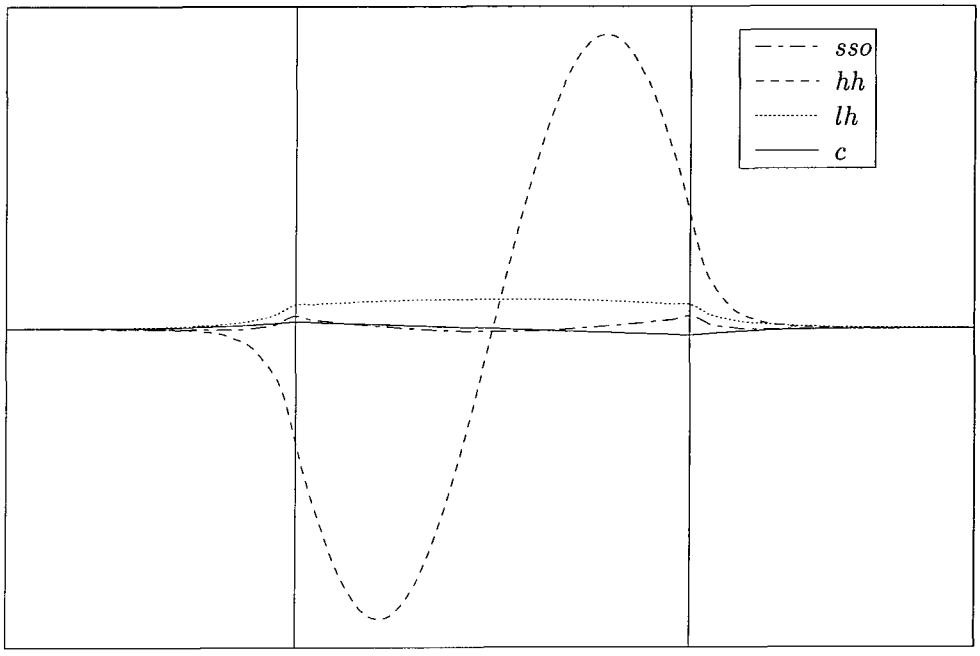


Figure 6.2 Dominant $f_n(z)$ for first excited electron bound state, at $\mathbf{k}_{||} = (0,0)$.



z direction

Figure 6.3 Dominant $f_n(z)$ for valence ground state, at $\mathbf{k}_{||} = (0,0)$.



z direction

Figure 6.4 Dominant $f_n(z)$ for first excited valence state, at $\mathbf{k}_{||} = (0,0)$.

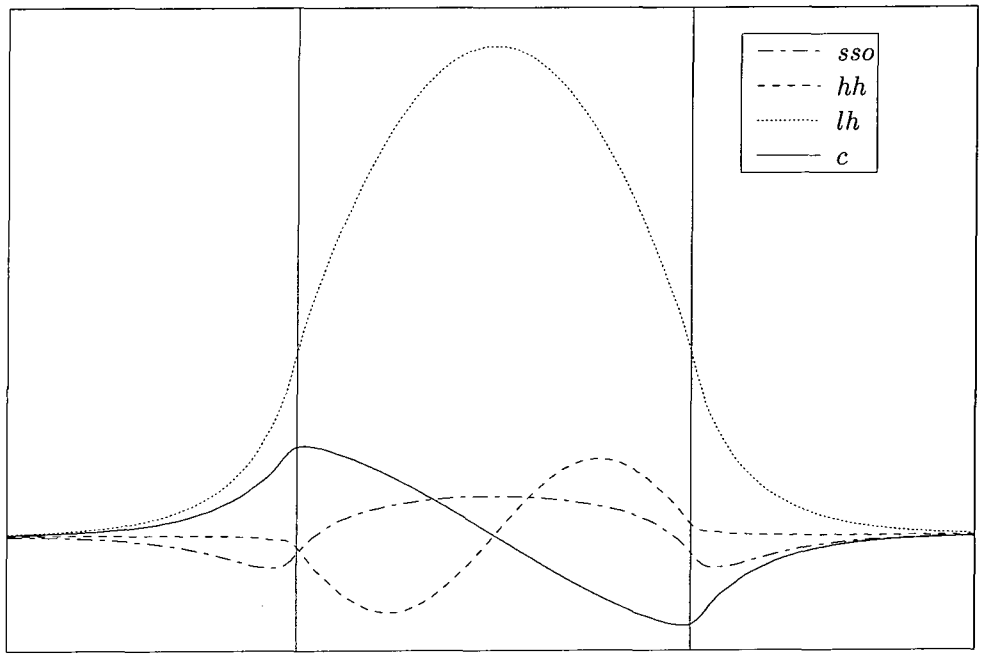


Figure 6.5 Dominant $f_n(z)$ for second excited valence state, which in this case is the light hole like ground state, at $\mathbf{k}_{||} = (0, 0)$.

6.2.2 Comparison to SPB model

To compare directly the results obtained from the EPM with an effective mass EFA, a simple one-dimensional single parabolic band (1D SPB) model is used. The relevant effective masses are obtained from the EPM calculations, rather than experiment, to ensure a fair comparison. This was done by direct measurement of the curvature of the relevant energy band, as calculated by the EPM. Recall that the effective mass tensor may be expressed in the form [111]

$$\frac{1}{m^*_{ij}} = \frac{1}{\hbar^2}\frac{\partial^2 E_{\mathbf{k}}}{\partial k_i \partial k_j}, \tag{6.1}$$

which simplifies to

$$\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{d^2 E_{\mathbf{k}}}{dk^2}, \tag{6.2}$$

as here we are only interested in the effective mass in the growth direction. The SPB quantum well depth is set to be the same as the band offsets used for the EPM calculation.

Since the SPB model has only one band, there is only one envelope function to consider, whilst there are usually several bands with non-negligible envelope functions derived from the EPM model. We concentrate here on the envelope function labelled c in Figure 6.1, which for conduction band ground states makes the dominant contribution to the envelope function expansion. This envelope function (and its spin degenerate 'twin') is multiplied by its associated zone centre Bloch periodic function for the first conduction band. This function and the one calculated using the SPB model are shown in Figure 6.6. Both have been normalised according to $\int_{\Omega_{sc}} f^* f dz = 1$. The two envelope functions are in very good agreement with each other.

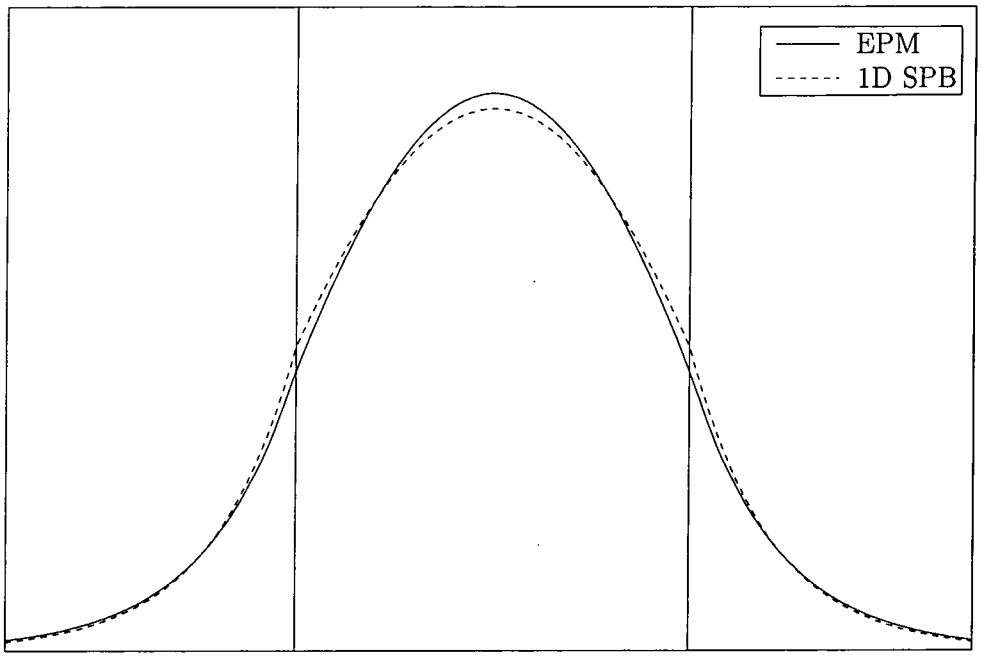


Figure 6.6 The dominant EPM-derived envelope function, $f_c(z)$, and the envelope as calculated using the 1D SPB model, for the conduction band ground state. Both envelopes have been normalised.

In order to examine the behaviour of the envelope functions at the interfaces, their spatial derivatives are calculated. However, the EPM-derived envelope function was first scaled so that it has the same value as the SPB envelope at the interfaces, to allow a more meaningful comparison between the two derivatives. The results are plotted in Figure 6.7.

The most obvious difference between the two results is that while the SPB derivative is discontinuous at the interfaces, the EPM derivative is continuous over all z. The discontinuity in the SPB gradient is a result of the boundary conditions applied at the interfaces (see Appendix C),

$$\frac{1}{m_b^*}\nabla_z f_b(z_0) = \frac{1}{m_w^*}\nabla_z f_w(z_0), \tag{6.3}$$

where m_b^* and m_w^* are the electron bulk effective masses in the barrier and well materials, respectively, and z_0 is the position of the interface. Thus, the ratio between

the discontinuous values is simply the ratio between the two effective masses, since the envelope, $f(z)$, is matched at the interface.

Burt has shown [17] how this boundary condition can be derived from the true boundary condition, that the derivatives of the full wavefunction match at the interfaces. Figure 6.7 shows that there is in fact a rapid change in the EPM-derived envelope function derivatives at the interfaces, and that at an interface this derivative is midway between the discontinuous values of the SPB model.

It is important to note that the EPM envelope functions are derived from a wavefunction that obeys the true boundary conditions, as obtained from the form of the (pseudo) Hamiltonian. Thus, this rapid change of envelope function derivative at interfaces is a genuine physical feature, and is due to the differences in the bulk properties of the two materials.

It should also be pointed out that the oscillations in the EPM envelope function, which are particularly apparent in the derivative, are a mathematical consequence of the particular envelope function theory adopted.

6.3 Dipole matrix elements

6.3.1 Background

The interaction between a charged particle and an electromagnetic radiation field can be expressed as either $\mathbf{A} \cdot \mathbf{p}$ or $\mathbf{E} \cdot \mathbf{r}$, depending on the choice of gauge (see Section 3.6). Either choice can be used without difficulty when dealing with heterostructure bound states, the $\mathbf{A} \cdot \mathbf{p}$ form involving the calculation of momentum matrix elements and the $\mathbf{E} \cdot \mathbf{r}$ form requiring the calculation of position, or *dipole*, matrix elements. However, some commonly used intuitive pictures and approximations lead to completely false predictions for the matrix elements of certain types of transitions

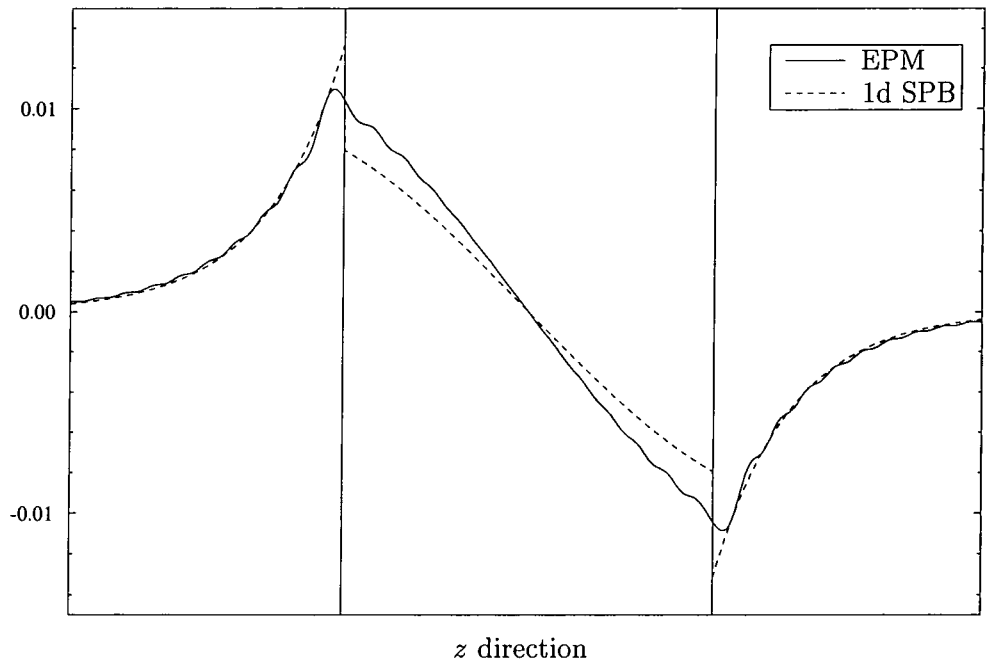


Figure 6.7 The gradients of the dominant EPM-derived envelope function and that calculated using the 1D SPB model, for the conduction band ground state.

when using the $\mathbf{E} \cdot \mathbf{r}$ form of the interaction, as has been pointed out by Burt [107,108]. Here, we are able to demonstrate the validity of Burt's predictions for real systems, by comparisons of dipole matrix elements calculated using the full wavefunction and approximations to it using the envelope function expansion method developed in the previous chapter.

6.3.2 Envelope function method

The calculation of the dipole matrix element between two quantum well bound states using the complex band structure method is discussed in Section 3.6. In this section, we consider calculating the dipole matrix element using the envelope function expansion of (5.5), namely

$$\Psi(\mathbf{r}) = \sum_n F_n(\mathbf{r}) u_n(\mathbf{r}). \tag{6.4}$$

In the supercell approach of Section 5.2.1, the wavefunction in all three regions of the system is described by a single Fourier representation. This simplifies the procedure for calculating the dipole matrix element compared to that described in Section 3.6, which can be written as

$$z_{nm} = \int_{z_0}^{z_1} \int_{\Omega_{||}} \left(\sum_n F_n(\mathbf{r}) u_n(\mathbf{r}) \right)^* z \sum_m F_m(\mathbf{r}) u_m(\mathbf{r}) \, d^2\mathbf{r}_{||} \, dz \tag{6.5}$$

$$= \sum_{nm} \int_{z_0}^{z_1} z f_n^*(z) f_m(z) \int_{\Omega_{||}} u_n^*(\mathbf{r}) u_m(\mathbf{r}) \, d^2\mathbf{r}_{||} \, dz, \tag{6.6}$$

where z_0 and z_1 are the end points of the supercell along the z-direction (the well growth direction). Again it is assumed that the two states, n and m, have the same in-plane wave vector and, as with the complex band structure calculation, the in-plane integration is over a unit cell face, $\Omega_{||}$.

Using the Fourier expansion for the Bloch periodic functions given in (5.9) it is possible for (6.6) to be rewritten in the form

$$z_{nm} = \int_{z_0}^{z_1} z f_{nm}(z) u_{nm}(z) \, dz, \tag{6.7}$$

where

$$f_{nm}(z) = f_n^*(z) f_m(z) \tag{6.8}$$

$$u_{nm}(z) = \sum_{\mathbf{g}\mathbf{g}'} u_{n\mathbf{g}}^* u_{m\mathbf{g}'} e^{i(g'-g)z} \, \Omega_{||} \, \delta_{\mathbf{g}_{||}\mathbf{g}'_{||}}, \tag{6.9}$$

and the z-dependent part of the envelope function, $f_n(z)$, is given by (5.13).

The accuracy of the reproduction of a quantum well wavefunction when performing only a partial sum over the envelope and Bloch periodic functions was investigated in Section 5.4. Two particular cases were examined, which we refer to

as the 'single band edge' and 'double band edge' approximations, or SBE and DBE,

$$\text{SBE:} \quad \Psi^{(j)} \approx F_j^{(j)}(\mathbf{r})u_j(\mathbf{r}) \tag{6.10}$$

$$\text{DBE:} \quad \Psi^{(j)} \approx F_v^{(j)}(\mathbf{r})u_v(\mathbf{r}) + F_c^{(j)}(\mathbf{r})u_c(\mathbf{r}) \tag{6.11}$$

where j indicates either a conduction band (c) or valence band (v) bound state. Thus, for example, a state bound in the conduction band quantum well has an SBE approximation including only the dominant term[3] which is the product of the conduction band Bloch periodic part and its corresponding envelope function,

$$\Psi^{(c)} \approx F_c^{(c)}(\mathbf{r})u_c(\mathbf{r}), \tag{6.12}$$

whilst the DBE approximation would also include the terms with the heavy hole, light hole, and spin split off u_n's. When it comes to predicting the probability density of a single bound state, the SBE approximation is very successful. In fact, the contribution from the dominant conduction band term to the total probability density of the conduction band bound state in the $55a/4$ wide $In_{0.53}Ga_{0.47}As/InP$ well is 97%. This leads one to expect that the SBE approximation should be a very useful one when predicting various electronic properties of quantum wells, including dipole matrix elements, and has been widely used as such (see for example [112–115]). However, Burt's algebraic approximations suggest otherwise, and it is the validity or otherwise of these approximations when calculating dipole matrix elements which is of particular interest here.

[3]In fact, due to spin degeneracy, the states come in pairs, so that the SBE approximation will actually have two terms (see Section 6.2.1).

	$\lvert z_{nm}\rvert$ (eÅ)		
Transition	EPM	DBE	SBE
c_1–c_2	15.793	15.751	15.431
c_1–hh_1	$5 \cdot 10^{-4}$	$2 \cdot 10^{-5}$	$4 \cdot 10^{-9}$
c_1–hh_2	0.636	0.634	$1 \cdot 10^{-8}$
c_1–lh_1	4.819	4.807	$1 \cdot 10^{-8}$
c_2–hh_1	0.002	0.002	$4 \cdot 10^{-8}$
c_2–hh_2	$5 \cdot 10^{-5}$	$6 \cdot 10^{-6}$	$1 \cdot 10^{-7}$
c_2–lh_1	0.010	$4 \cdot 10^{-4}$	$3 \cdot 10^{-7}$
hh_1–hh_2	11.766	11.688	11.688
hh_1–lh_1	1.408	1.392	1.393

Table 6.1 Magnitude of the dipole matrix elements for an $In_{0.53}Ga_{0.47}As/InP$ single quantum well of width $55a/4 = 80.696$Å, calculated using the full complex band structure EPM and the double and single band edge approximations to the envelope function expansion.

6.3.3 Results

Table 6.1 lists the magnitudes of the (generally complex) dipole matrix elements between pairs of states in the $55a/4$ wide $In_{0.53}Ga_{0.47}As/InP$ single quantum well, as calculated using the complex band structure method, DBE and SBE approximations. The dipole matrix element between the same state is merely the position expectation value of that state, i.e., $z_{nn} = \langle z \rangle$, for state n, hence this must be real. In this system, the potential is symmetric about the centre of the well,[4] and so by setting the z origin to be in the centre of the well, the expectation value is zero. In theory, the states in the quantum well are orthonormal, so the position of the z-origin origin should have no effect when calculating the dipole matrix element between states [106]. However, in practice some pairs of different states are not truly orthogonal,[5] due to the use of

[4]Strictly speaking, only the in-plane averaged potential is symmetric since the zinc-blende structure does not have inversion symmetry along the primary axes.

[5]Most pairs of states *are* orthonormal to within machine precision, and so the z origin really has no effect. The worst case for this system was a 0.6% overlap between the hh_2 and lh_1 states. Whilst this is a small error in terms of the overlap, any finite overlap between states leads to a dipole moment that is proportional to the z origin, so the apparent error in the dipole moment

a finite number of reciprocal lattice vectors in the original complex band structure calculation, hence it is important to use the same origin for all methods used in calculating the dipole moment to ensure a fair comparison.

Comparison in Table 6.1 of the DBE results to those using the full wavefunctions in the calculations shows that the dipole matrix elements predicted by the DBE approximation are at least qualitatively reliable, and generally numerically reliable as well. The matrix elements predicted by the DBE approximation tend to be slightly too small, but it must be remembered that terms that make small but finite contributions to the full wavefunctions have been dropped in the envelope function expansions, hence this is to be expected. Thus, at least for dipole matrix element calculations, the DBE approximation is a perfectly valid one.

We now turn our attention to the SBE approximation. Table 6.1 shows that for the *intra*band transitions considered (c_1–c_2, hh_1–hh_2 and hh_1–lh_1), the SBE is as successful as the DBE approximation in predicting dipole matrix elements. However, the SBE approximation fails completely to predict the correct *inter*band dipole matrix elements [107,108]. This is not just a quantitative failure due to the somewhat cruder approximation to the full wavefunction of the SBE compared to DBE envelope function expansion, but a qualitative failure as well. Consider the dipole matrix element between the ground conduction and light hole states (c_1–lh_1), corresponding to the fundamental strong TM mode transition. Both the full EPM and the DBE approximation predict a dipole matrix element that is equivalent to approximately 6% of the well width. In stark contrast,the SBE approximation predicts a dipole matrix element equivalent to only 10^{-8}% of the well width!

could be large.

6.3.4 Discussion

The dipole matrix elements for intraband transitions are as expected, i.e., they are on the same scale as the well width — equivalent to as much as 20% of the well width for the c_1–c_2 transition. This is the expected result, since the envelope functions of the ground and first excited states have s and p-like symmetry with respect to the centre of the well, and the Bloch periodic state multiplying this is the same in both cases, $u_c(\mathbf{r})$. The envelope functions vary on the scale of the well, and hence we expect dipole matrix elements on the same scale. Thus, there is little difference between the predictions of the SBE and DBE approximations and the full wavefunction calculation.

The results for interband transitions, however, are initially somewhat surprising. Since the probability density of a state can be reproduced to within a few percent by the SBE approximation, one would expect calculations of dipole matrix elements using such an approximation to be at least qualitatively correct, as is indeed the case for intraband transitions. For interband transitions, the SBE approximations for the wavefunctions of the two states are

$$\Psi^{(c)}(\mathbf{r}) = F_c^{(c)}(\mathbf{r})u_c(\mathbf{r}) \tag{6.13}$$

$$\Psi^{(v)}(\mathbf{r}) = F_v^{(v)}(\mathbf{r})u_v(\mathbf{r}). \tag{6.14}$$

Thus, using the same reasoning as for the intraband case, one would expect dipole matrix elements on the atomic scale for interband transitions between ground states, since both states have an s-like envelope function with the only significant differences being between the Bloch periodic parts, $u_c(\mathbf{r})$ and $u_v(\mathbf{r})$. As these vary within the primitive unit cell, the 'atomic' scale, but are periodic with the lattice, one would expect dipole matrix elements on the atomic scale. This is not the case, and interband dipole matrix elements increase with the well width. In fact, calculations

by the author predict a 13.6eÅ c_1–lh_1 dipole matrix element for a 153.9Å wide InSb/CdTe quantum well, which is certainly not on the atomic scale. Obviously, the SBE approximation has omitted detail that is vital to the correct evaluation of the interband dipole matrix elements. Burt [108] has shown that this is the contribution of the 'cross' terms,

$$\left\langle F_v^{(c)} u_v \middle| z \middle| F_v^{(v)} u_v \right\rangle + \left\langle F_c^{(c)} u_c \middle| z \middle| F_c^{(v)} u_c \right\rangle, \tag{6.15}$$

which dominate the interband matrix element, even though they derive from terms which play only a minor role in the envelope function representation of the wavefunction. These terms have similar features to those of intraband matrix elements in the SBE approximation, i.e., a common Bloch periodic function and envelopes of differing parity. Thus, large dipole matrix elements are obtained, though these are not as large as for intraband transitions since $F_v^{(c)}$ and $F_c^{(v)}$ are of smaller magnitude than $F_c^{(c)}$ and $F_v^{(v)}$, as can be seen from Figures 6.1 and 6.5.

6.4 Charge oscillation

6.4.1 Background

We now turn our attention to the charge oscillation induced by coherent interband optical excitation. The starting point is the time-dependent Schrödinger equation,

$$\left[-\frac{\hbar^2}{2m_e} \nabla^2 + V(\mathbf{r}) \right] \Phi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t} \Phi(\mathbf{r}, t). \tag{6.16}$$

The solutions take the usual form

$$\Phi(\mathbf{r},t) = \Psi(\mathbf{r})\, e^{-i\omega t} \tag{6.17}$$

$$\omega = \frac{E}{\hbar}, \tag{6.18}$$

where E is the total energy of the state and $\Psi(\mathbf{r})$ is the solution to the time-independent Schrödinger equation.

Suppose at time $t = 0$ the electron is in the state

$$\Phi(\mathbf{r},0) = \frac{1}{\sqrt{2}}\left(\Phi^{(v)}(\mathbf{r},0) + \Phi^{(c)}(\mathbf{r},0)\right) \tag{6.19}$$

$$= \frac{1}{\sqrt{2}}\left(\Psi^{(v)}(\mathbf{r}) + \Psi^{(c)}(\mathbf{r})\right), \tag{6.20}$$

which is the special case of a 1:1 combination of valence and conduction band states. The equal weight is chosen to emphasise the results by producing the maximum amount of charge oscillation. If the electron is coherently excited by a laser, then at a later time, t, it will be in the state

$$\Phi(\mathbf{r},t) = \frac{1}{\sqrt{2}}\left(\Psi^{(v)}(\mathbf{r})\, e^{i\omega_v t} + \Psi^{(c)}(\mathbf{r})\, e^{i\omega_c t}\right), \tag{6.21}$$

so that the charge density will be (in units of e, with c.c. indicating the complex conjugate term)

$$|\Phi(\mathbf{r},t)|^2 = \frac{1}{2}\left(\left|\Psi^{(v)}\right|^2 + \left|\Psi^{(c)}\right|^2 + \left(\Psi^{(v)}\right)^* \Psi^{(c)}\, e^{-i\omega_{cv}t} + \text{c.c.}\right) \tag{6.22}$$

with the period of oscillation, T, determined thus,

$$\omega_{cv} = \omega_c - \omega_v \tag{6.23}$$

$$= (E_c - E_v)/\hbar \tag{6.24}$$

$$= \frac{2\pi}{T}. \tag{6.25}$$

Thus, the charge density is made up of two terms that are static in time,

$$\frac{1}{2}\left(\left|\Psi^{(v)}\right|^2 + \left|\Psi^{(c)}\right|^2\right), \tag{6.26}$$

and two oscillatory terms, labelled as $\mu(\mathbf{r}, t)$,

$$\mu(\mathbf{r}, t) = \frac{1}{2}\left(\left(\Psi^{(v)}\right)^* \Psi^{(c)} e^{-i\omega_{cv}t} + \text{c.c.}\right). \tag{6.27}$$

6.4.2 Results

The results in this section are for a InSb/CdTe single quantum well of width $55a/4$, where $a = 6.48$Å. Since the conduction band does not couple strongly with the heavy hole band for an EM electric field parallel to the growth direction (TM mode), we consider excitation between the light hole and conduction band ground states.

Figure 6.8 shows the in-plane averaged charge (i.e., probability) density, calculated directly using the complex band structure method. One can clearly see that the charge oscillates on the scale of the well width. This is interesting because it has been pointed out [107] that the charge might intuitively be expected to oscillate only on the atomic scale, rather than what could be termed the 'quantum well' scale. This intuitive expectation is based on similar assumptions to those discussed in Section 6.3.3 with regard to the calculation of interband dipole matrix elements, and is because the probability density of the bound states can be reasonably approx-

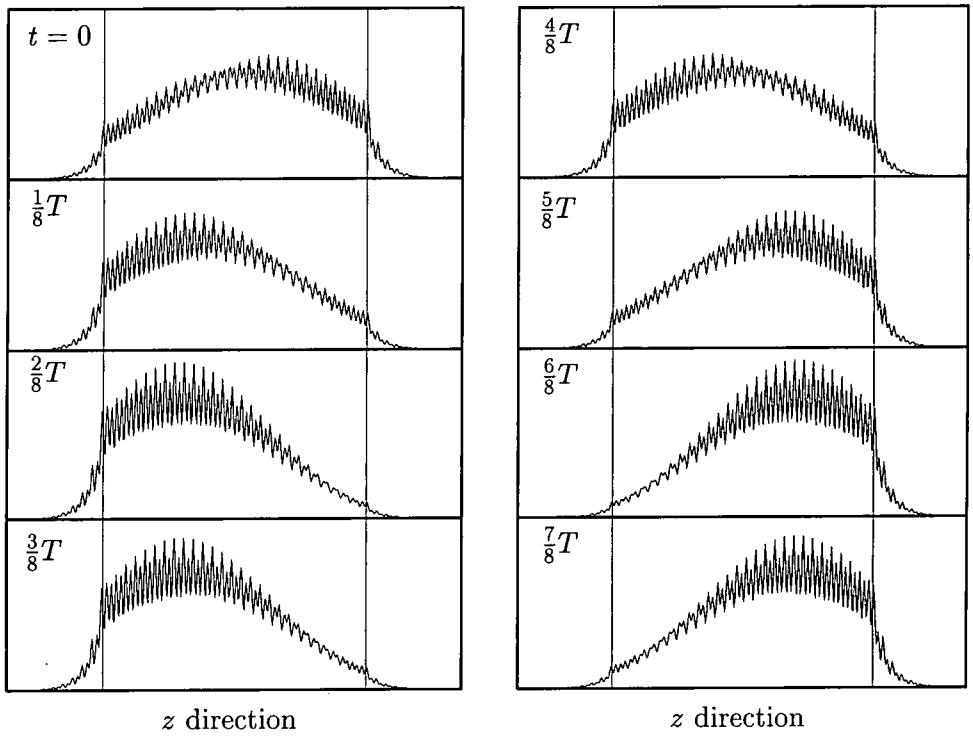


Figure 6.8 The in-plane averaged charge density, $|\widetilde{\Phi}|^2$, for a $55a/4$ InSb/CdTe well, as calculated using the complex band structure EPM approach. Such charge oscillation could occur due to coherent interband optical excitation. The charge density is shown at time intervals of $T/8$.

imated by envelope function expansions including only the appropriate SBE terms, as demonstrated in Section 5.4. Conduction and valence ground states have very similar envelope functions but differ in the nature of their Bloch periodic parts, one of which will be that for the valence band edge and the other for the conduction band edge. These characterise the atomic aspects of the crystal, and the time dependence of a simple linear combination of just these dominant terms would give oscillations on the atomic length scale.

Figure 6.9 shows the DBE approximation to the charge density, as in Figure 6.8. As with the dipole matrix element calculations, the DBE approximation faithfully reproduces the result obtained with the full complex band structure calculation, for the c_1–lh_1 case at least. In contrast, the SBE approximation to the charge density, plotted in Figure 6.10, predicts negligible charge oscillation on the quantum well scale, but merely oscillations on the atomic scale. It is apparent that, as in the dipole moment, the 'other' band edge terms play a critical role in describing the charge oscillation. Again, it is the 'cross' terms in the charge density, which are included in the DBE but missing from the SBE approximations, that are responsible for the large-scale oscillations in Figures 6.8 and 6.9.

The DBE approximations for the spatial part of the valence and conduction state wavefunctions are[6]

$$\Psi^{(v)} \approx F_v^{(v)} u_v + F_c^{(v)} u_c \tag{6.28}$$

$$\Psi^{(c)} \approx F_v^{(c)} u_v + F_c^{(c)} u_c. \tag{6.29}$$

Therefore, the spatial part of the first term in the charge density modulation term,

[6]The spin degeneracy of the bands and the sum over the light and heavy hole bands is made implicit here, for notational simplicity.

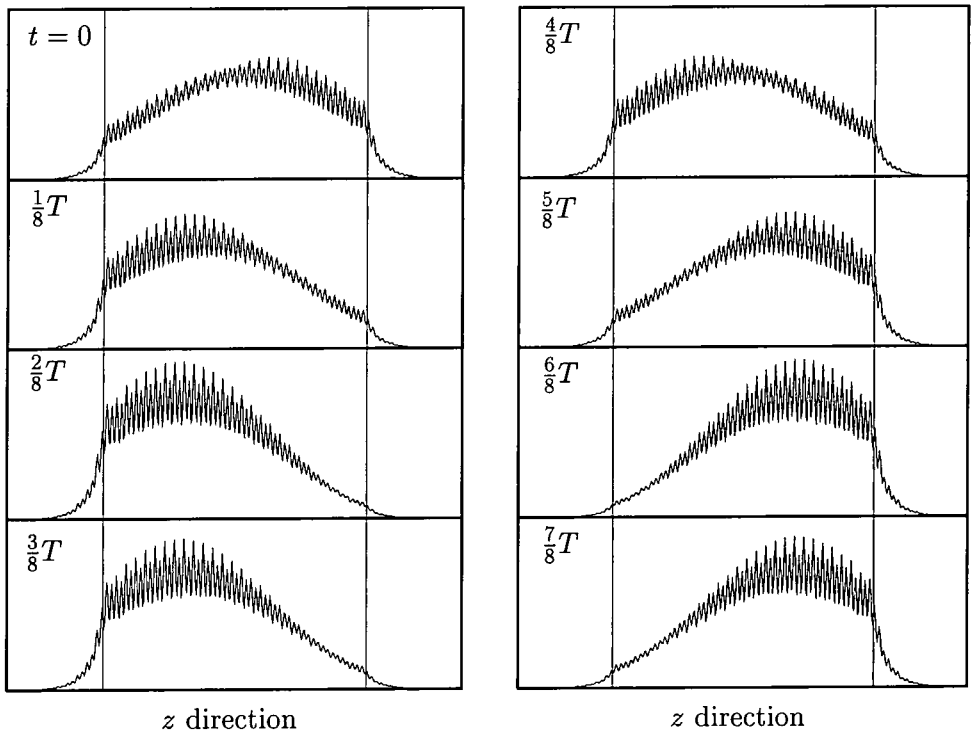


Figure 6.9 The in-plane averaged charge density, $|\widetilde{\Phi}|^2$, for a $55a/4$ InSb/CdTe well, as calculated using the DBE approximation. Such charge oscillation could occur due to coherent interband optical excitation. The charge density is shown at time intervals of $T/8$.

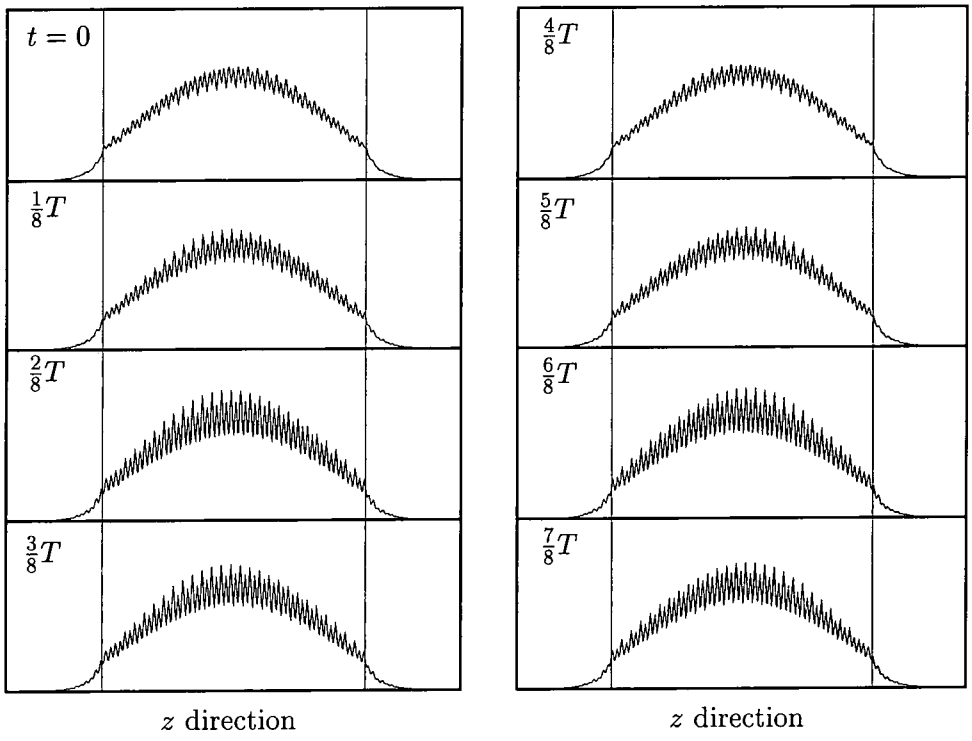


Figure 6.10 The in-plane averaged charge density, $|\widetilde{\Phi}|^2$, for a $55a/4$ InSb/CdTe well, as calculated using the SBE approximation. Such charge oscillation could occur due to coherent interband optical excitation. The charge density is shown at time intervals of $T/8$.

$\mu(\mathbf{r},t)$, is

$$\begin{aligned}\left(\Psi^{(v)}\right)^* \Psi^{(c)} \approx \left(F_v^{(v)}\right)^* F_v^{(c)} |u_v|^2 + \left(F_c^{(v)}\right)^* F_c^{(c)} |u_c|^2 \\ + \left(F_v^{(v)}\right)^* F_c^{(c)} u_v^* u_c + \left(F_c^{(v)}\right)^* F_v^{(c)} u_c^* u_v.\end{aligned} \tag{6.30}$$

The states under consideration here are the c_1 and lh_1 bound states, whose dominant envelope functions are plotted in Figures 6.1 and 6.5. These show that $F_{lh}^{(lh)}$ and $F_c^{(lh)}$ have opposite parity, with respect to the centre of the well, as do $F_{lh}^{(c)}$ and $F_c^{(c)}$. Thus, their products have odd parity leading to the spatial variation in $\mu(\mathbf{r},t)$ required for the charge oscillation to be on the scale of the quantum well. In the SBE approximation, (6.30) reduces to

$$\left(\Psi^{(v)}\right)^* \Psi^{(c)} \approx \left(F_v^{(v)}\right)^* F_c^{(c)} u_v^* u_c, \tag{6.31}$$

in which the envelope functions both have even parity with respect to the centre of the well, so that their product also has even parity. Thus, using the SBE $\mu(\mathbf{r},t)$ cannot produce charge oscillations over the whole of the well, but only over the atomic scale, due to the factor of $u_v^* u_c$, which is periodic with the lattice.

6.5 Operator ordering in EFAs

This section contains work that is the result of a collaboration between the author and Paul Stavrinou, of the Interdisciplinary Research Centre for Semiconductor Materials, Imperial College of Science, Technology and Medicine.

6.5.1 Introduction

Section 6.2 illustrates that it is possible to express the electron wavefunction in terms of envelope functions that are uniquely defined, and are continuous in value and

gradient, even across heterojunctions, as predicted by Burt [17]. Figure 6.7 confirms that the discontinuity in the gradient of an envelope function at an interface, which has been obtained using an approximate rather than exact theory, arises as an approximation to the smooth but rapid change in the gradient of the exact envelope function. It also supports the traditional choice of boundary conditions, the continuity of $f(z)$ and $(m^*)^{-1}\nabla_z f(z)$ (see Appendix C). However, the comparison between exact and approximate envelope functions used only a single parabolic band for the latter. In reality, the commonest EFAs used are based on more sophisticated $\mathbf{k}\cdot\mathbf{p}$ models, using a number of interacting bands instead of a single isolated one.

The usual, well established approach to using $\mathbf{k}\cdot\mathbf{p}$ perturbation theory when modelling heterostructures is to start with the Kane [95] form for the matrix elements of the Hamiltonian for bulk crystals. Kane partitions the electron states into two categories, here labelled A and B. Category A contains the states corresponding to the bands of interest and those that interact strongly with them, whilst category B contains states that interact only weakly with those in A. This allows the use of the perturbation scheme developed by Löwdin [116], in which the weak interactions between states in A and B are removed while interactions between the states in A are 'renormalised'. The matrix containing the renormalised A states is then diagonalised exactly. As Kane showed, this scheme lends itself well to the description of the bands of a bulk semiconductor close to the zone centre and around the fundamental band gap, i.e., the light and heavy hole, spin split off and lowest conduction bands (labelled *lh*, *hh*, *sso* and *c*, respectively). These are the natural choice for the category A states, with the remaining, energetically remote bands assigned to category B and included in the Hamiltonian only as perturbations.

The usefulness of this method comes from the fact that the perturbations due to the states in B can be included using a small number of parameters; the effective masses and energy separations of the bands in A. These masses and energies may be

readily determined experimentally. If one is interested in low energy valence band states only, then the conduction band and spin split off states may be reassigned to B, depending on the strength of their interaction with, and energy separation from the light and heavy hole states.

Kane's model has proved very successful for bulk materials. However, its application to heterostructures has, until recently, not been without its problems. These problems arise because the Kane model was developed for bulk homogeneous crystals, and so the form of the Hamiltonian at abrupt heterojunctions was a matter of some debate. This is because, for calculating heterostructure states, it is necessary to include the variation of system properties in the growth direction. At first sight, this can be achieved by replacing k_z with $-i\partial/\partial z$ in the original Kane Hamiltonian [92,117]. However, if a straight substitution of the type

$$k_z \mapsto -i\frac{\partial}{\partial z} \tag{6.32}$$

$$k_z^2 \mapsto -\frac{\partial^2}{\partial z^2} \tag{6.33}$$

is made, the resultant Hamiltonian is non-Hermitian, and so, to ensure Hermiticity (and also conservation of current density across a heterojunction), some authors have 'symmetrised' the Hamiltonian [92,117]. The problem[7] with this approach is that there are an infinite number of possible symmetrisations [118] (cf. Section C.1, which discusses this in the context of the SPB approximation).

This ambiguity was resolved as a consequence of Burt's derivation of effective mass envelope function equations from first principles [17]. The derivation shows for inhomogeneous systems, including heterostructures, how the remote bands should

[7]There were a number of problems with the traditional $\mathbf{k}\cdot\mathbf{p}$ method for heterostructure calculations. One other notable inconsistency is the assumption made during the calculation of heterostructure states that the Bloch periodic functions, $u_n(\mathbf{r})$, do not change between materials, even though the elements of the Kane-like Hamiltonian used for each material include the correct material dependence.

be included in the Schrödinger-like approximate envelope function equation. This equation has been used by Foreman [16] to calculate a new, so called 'ordered', $\mathbf{k} \cdot \mathbf{p}$ Hamiltonian, H_{ord}, applicable to heterostructures, without the need for the traditional symmetrisation procedure. Foreman [16] has shown that the traditional symmetrised Hamiltonian [92,117], H_{sym}, results in incorrect matching conditions, in the common approach of obtaining general solutions for the wavefunction on each side of a heterojunction and then matching them at the interface. Further, if the conduction band is included as a perturbation (category B), calculations for valence band quantum well bound states using the two Hamiltonians sometimes predict substantially differing effective masses. This has been further demonstrated by Stavrinou and van Dalen [99], who calculated in-plane band structure using the the two Hamiltonians.

The purpose of this section is to demonstate that $\mathbf{k} \cdot \mathbf{p}$ heterostructure calculations for quantum wells which use the ordered Hamiltonian give in-plane band structure in agreement with pseudopotential calculations, whereas that is not always the case with the symmetrised Hamiltonian.

6.5.2 Calculations

The elements of the ordered Hamiltonian, H_{ord}, that differ from the symmetrised Hamiltonian, H_{sym}, arise naturally from Burt's first principles derivation of the effective mass envelope function equation, which, for band a, is [17]

$$-\frac{\hbar^2}{2m_e}\sum_{a'}\nabla\cdot(\gamma_{aa'}\cdot\nabla F_{a'}(\mathbf{r}))+H^{(2)}_{aa}(E,\mathbf{r})F_a(\mathbf{r})=EF_a(\mathbf{r}), \tag{6.34}$$

where $\gamma_{aa'}$ is the effective mass tensor, $H^{(2)}_{aa}$ approximates the a^{th} band edge profile, due to a band offset, and the remaining symbols have their usual meaning. The

effects of the remote bands are included in $\gamma_{aa'}$, which takes the form

$$\gamma_{aa'}(E,\mathbf{r}) = \delta_{aa'} + \frac{2}{m_e}\sum_b \frac{p_{ab}p_{ba'}}{E - E_b(\mathbf{r})}, \tag{6.35}$$

where the sum is over all remote bands, b, p_{ab} is the momentum matrix element between a and b, and $E_b(\mathbf{r})$ is the energy of band b, which depends on the material, and therefore the position in the heterostructure. The states labelled a/a' and b are in categories A and B, respectively. The important point to note with the first (kinetic energy) term in (6.34) is that the ordering of the differential operators is not the result of any assumptions or symmetrisation procedures, but emerges when the equation is derived from the full Schrödinger equation.

In this section, we include only the light, heavy hole and spin split off bands explicitly, i.e., the category A states are *lh*, *hh* and *sso*, with all remaining bands being in B. Following Foreman [16], we replace the effective mass tensor with three dimensionless parameters that describe the interactions between the p-type valence band zone centre states, $|X\rangle$, $|Y\rangle$ and $|Z\rangle$, and the remote bands, b,

$$\sigma = \frac{1}{3m_e}\sum_b^{\Gamma_1} \frac{\left|\langle X|\hat{p}_x|b\rangle\right|^2}{E_b - E_v} \tag{6.36a}$$

$$\pi = \frac{1}{3m_e}\sum_b^{\Gamma_{15}} \frac{\left|\langle X|\hat{p}_y|b\rangle\right|^2}{E_b - E_v} \tag{6.36b}$$

$$\delta = \frac{1}{6m_e}\sum_b^{\Gamma_{12}} \frac{\left|\langle X|\hat{p}_x|b\rangle\right|^2}{E_b - E_v}, \tag{6.36c}$$

where each summation is over states of the same symmetry type,[8] E_v is the valence band edge energy in the absence of spin orbit coupling, and (without loss of generality) we only consider finite momentum matrix elements between the remote states

[8]Only states with symmetries corresponding to s, p and d orbitals are included since the effect on the valence band of other orbitals is negligible.

Material	σ	π	δ	Δ_0	ΔE_v
$In_{0.53}Ga_{0.47}As$	3.236	0.512	0.140	0.370	0.38
InP	1.939	0.488	0.138	0.110	0
InSb	4.658	0.419	0.136	0.980	0.25
CdTe	0.990	0.306	0.127	0.840	0

Table 6.2 The EPM derived valence band $\mathbf{k} \cdot \mathbf{p}$ parameters for the $In_{0.53}Ga_{0.47}As$/InP and ersatz InSb/CdTe quantum wells. The parameters σ, π and δ are defined in (6.36). Δ_0 and Δ_v are the spin split off energy and valence band offset, respectively. Note that the InSb/CdTe valence band offset has been artificially reduced from its true value of 0.87eV. The reason for this is explained in the text. These parameters were used by Paul Stavrinou to obtain the $\mathbf{k} \cdot \mathbf{p}$ quantum well band structure plotted in Figures 6.11 and 6.12.

and $|X\rangle$. The symmetry types are those applicable in the absence of spin orbit coupling, which is included in the Hamiltonian at a later stage. These three parameters may be expressed in terms of the more common Luttinger parameters [119],

$$\gamma_1 = -1 + 2\sigma + 4\pi + 4\delta \tag{6.37a}$$

$$\gamma_2 = \sigma - \pi + 2\delta \tag{6.37b}$$

$$\gamma_3 = \sigma + \pi - \delta. \tag{6.37c}$$

For the work presented in this section, the author calculated values for the σ, π and δ parameters in (6.36) within the framework of the EPM, for each material in the quantum wells investigated. This was done with the spin orbit coupling 'switched off' (i.e., S_μ was set to zero, cf. Section 2.2.3) and care was taken to ensure that the valence band state used in the calculations had the correct X-type symmetry. These values, along with the band offsets and spin orbit splittings, for the $In_{0.53}Ga_{0.47}As$/InP and ersatz InSb/CdTe systems are given in Table 6.2.

6.5.3 Results

The in-plane energy band structure for a $55a/4$ wide $In_{0.53}Ga_{0.47}As/InP$ valence band quantum well is shown in Figure 6.11. The figure compares band structure calculated by the author using the full complex band structure EPM method (see Chapter 3) and that calculated by Paul Stavrinou, using the symmetrised and ordered $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians. Unfortunately, these results are not very useful for the purpose of demonstrating the validity of the ordered over symmetrised Hamiltonians, since the difference between the bands calculated by the two Hamiltonians is small and on the scale of the errors, with respect to the EPM data, inherent in using a $\mathbf{k} \cdot \mathbf{p}$ method that explicitly includes the top valence bands only. The figure does at least provide reassurance that for this system the three ($\times 2$, for spin) band $\mathbf{k} \cdot \mathbf{p}$ method is in very good agreement with EPM predictions, over the energy range shown.

The small difference between the band structures of the symmetrised and ordered Hamiltonians is predominantly due to the small ratio between $In_{0.53}Ga_{0.47}As$ and InP σ parameters (see [16] for a discussion of the importance of σ). To obtain large differences in the band structure generated from the two Hamiltonians, we require materials with a greater ratio between their σ's. In addition, we note that narrower wells should also enhance the differences, since we require appreciable barrier penetration by the envelope functions in order for changes in the boundary conditions, due to differences in the Hamiltonians, to be noticeable. For these reasons, calculations were performed for a second system — a $12a/4$ wide InSb/CdTe quantum well.

Table 6.2 shows that the $\sigma^{(\text{well})} : \sigma^{(\text{barrier})}$ ratio for the InSb/CdTe material system is approximately 4.7:1, in contrast to a ratio of 1.7:1 for the $In_{0.53}Ga_{0.47}As/InP$ system. This suggests that the band structures calculated by the two $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians should differ in a far more noticeable way than for $In_{0.53}Ga_{0.47}As/InP$. Indeed

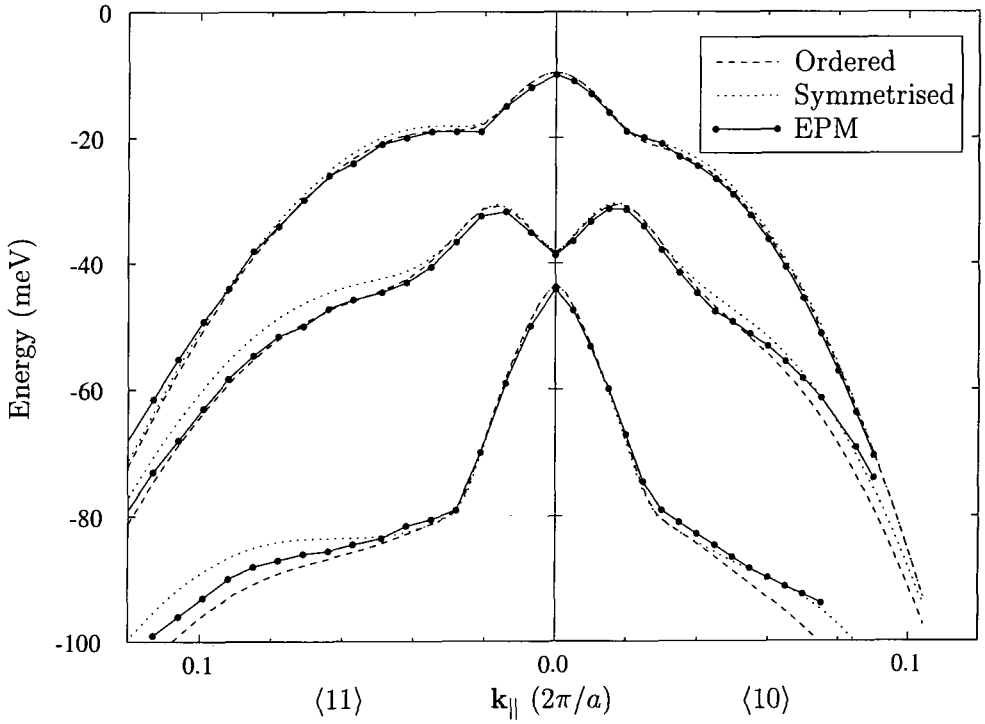


Figure 6.11 Valence energy band structure for a $In_{0.53}Ga_{0.47}As/InP$ quantum well, of width $55a/4$. The figure shows the comparison between the band structure predicted by the full EPM calculation and that predicted by the $\mathbf{k} \cdot \mathbf{p}$ method, using both ordered and symmetrised Hamiltonians. The $\mathbf{k} \cdot \mathbf{p}$ data was calculated by Paul Stavrinou, using bulk material parameters provided by the author. The EPM data is the average of the spin split bands.

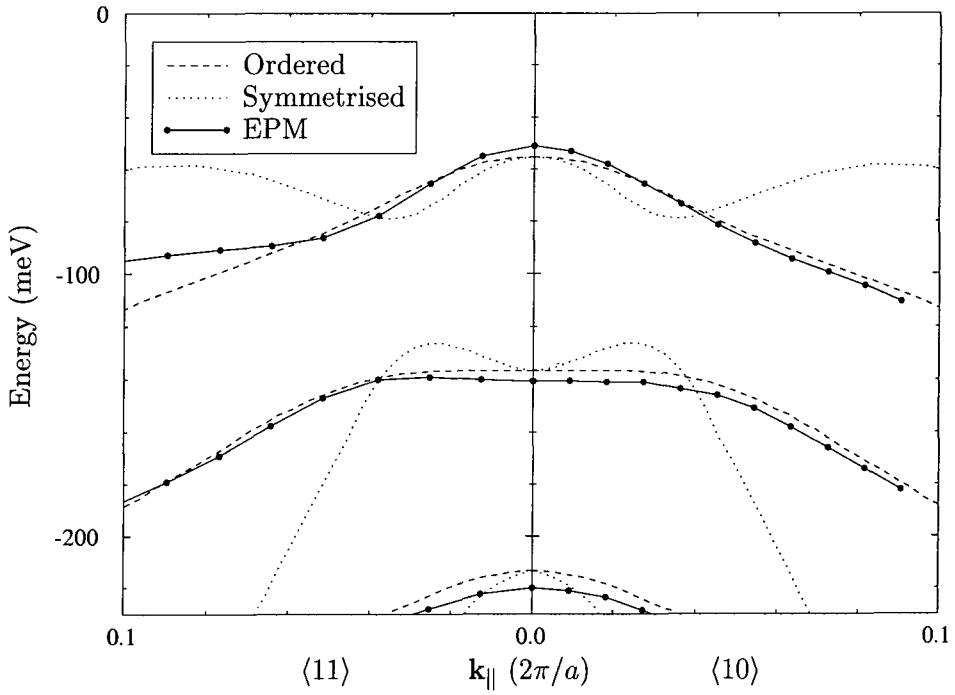


Figure 6.12 Valence energy band structure for the ersatz InSb/CdTe quantum well, of width $12a/4$ and with the reduced VB offsetof 0.25eV. The figure shows the comparison between the band structure predicted by the full EPM calculation and that predicted by the $\mathbf{k} \cdot \mathbf{p}$ method, using both ordered and symmetrised Hamiltonians. The $\mathbf{k} \cdot \mathbf{p}$ data was calculated by Paul Stavrinou, using bulk material parameters provided by the author.The EPM data is the average of the spin split bands.

this is the case, but there are a different set of problems associated with this system. Firstly, the correct well depth, 0.87eV, is so deep that even the ground and first few excited bound states occur at energies and wave vectors sufficiently far away from the valence band edge that the bulk $\mathbf{k} \cdot \mathbf{p}$ method is becoming increasingly inaccurate compared to the EPM. Secondly, the spin orbit coupling in InSb is particularly strong, leading to a spin split off energy of 0.98eV, compared to 0.37eV for $In_{0.53}Ga_{0.47}As$, whilst the fundamental band gap is small, 0.18eV. Unfortunately, the $\mathbf{k} \cdot \mathbf{p}$ method employed does not deal with this extreme situation very well. However, by reducing the InSb/CdTe valence band offset to 0.25eV to create an ersatz system, some very satisfactory results are obtained.

Figure 6.12 shows the valence band structure for the ersatz InSb/CdTe quantum

well, of width $12a/4$. The difference between dispersion curves of the symmetrised and ordered Hamiltonians is quite striking. We also note that the energy bands obtained using the ordered Hamiltonian are in excellent agreement with those of the full EPM calculation, in contrast to those calculated using the traditional, symmetrised Hamiltonian. We consider this to be conclusive proof of the validity of the ordered Hamiltonian and the incorrectness of the traditional symmetrised Hamiltonian.

Chapter 7

Summary and conclusion

This thesis presents theoretical methods and results for the study of electronic states in semiconductor heterostructures. The foundation for the work is the empirical pseudopotential method (EPM).

In Chapter 2, a description of the EPM is given, concentrating on the aspects and modifications necessary for heterostructure calculations. This demonstrates the solid physical basis of the method, highlighting its genuinely microscopic nature, i.e., the resolution of individual atoms. The chapter describes how the important effects of spin orbit coupling and strain are readily introduced. The chapter ends with typical examples of the bulk crystal properties of InP; full zone energy band structure, (single) density of states, imaginary part of the dielectric function and valence charge density. These all illustrate that the EPM can give an accurate description of energy bands and wavefunctions throughout the Brillouin zone. Such bulk calculations are important as a means of checking that the empirically determined pseudopotential parameters provide not only a good description of the particular data to which they were fitted, but also of the electronic properties as a whole. This is especially useful when, as in this work, the parameter fitting was done to a limited number of transition energies at high symmetry points in the Brillouin zone. That the

calculated density of states, dielectric function and real space charge density all produce expected results is a great source of comfort before embarking on heterostructure calculations.

Chapter 3 describes the complex band structure EPM method used to calculate the energies and wavefunctions of bound states in heterostructures of arbitrary configuration. The only restriction imposed is that the states under consideration must be genuinely bound, i.e., there must be no undecaying propagating states in the extreme left and right barrier regions.

Knowledge of the wavefunctions enables the calculation of both the momentum and position (or dipole) matrix elements between bound states, and hence an estimate of the relative transition rates, induced by electromagnetic (EM) radiation, of pairs of states. Generation of the in-plane energy band structure for pairs of states allows their joint density of states to be calculated. This also effects the rates of EM induced transitions. The methods for calculating these quantities are described in the chapter, including the approximations required to make the calculation of the joint density of states practicable. At the end of the chapter is a description of the method used to approximate the effect of applying an electrostatic field along the heterostructure growth direction.

An important point about the methods described in Chapter 3 is their flexibility. Valence and conduction band offsets can be adjusted down to monolayer precision. This was originally built into the software by the author to allow for the modelling of heterostructures consisting of an arbitrary choice of materials. However, this feature could also be used to approximate band bending effects, for example. It should also be noted that, although not utilised for the work presented here, the software allows for the effects of strain and an electrostatic field, the in-plane band structure, probability (hence charge) density and momentum matrix elements to all be calculated for any structure, provided the states of interest are bound.

The power and versatility of the methods described in Chapter 3 are, at least partially, illustrated in Chapter 4. The bound state energies of InAs/AlSb quantum wells are in good agreement with theoretical and experimental data of other workers, even though no effort had been made to maximise the agreement. This suggests that were one interested in modelling a specific system as accurately as possible, the appropriate tailoring of the input data to the EPM parameter fitting procedures should result in a heterostructure calculation capability that would be a very powerful predictive and analytic tool.

The predictions for the InAs/AlSb system are that there is the strong possibility of EM induced transitions within the conduction band at the technologically very important wavelength of 1.55μm. Because the transition is an intraband one, the joint density of states is much larger than for interband transitions, and we also expect an enhanced dipole (hence momentum) matrix element compared to transitions between electron and hole ground states. This promising narrow well system merits further investigation in terms of specific device applications.

The second part of Chapter 4 investigates the behaviour of the InAs/GaSb/AlSb asymmetric well bound states in the region of the InAs/GaSb band overlap. It was found that hybridization and the associated band anti-crossing occurs between electron-like and light hole-like states as a function of both the InAs layer width and the applied electric field strength. This is analogous to the appearance of minigaps in the in-plane band structure of InAs/GaSb heterostructures, reported by other workers, except that there is no electron-heavy hole interaction, since the coupling between these bands is negligible at the 2D zone centre. The anti-crossing between some of the states modifies the response of their energies to an applied electrostatic field from the strictly linear variation normally obtained in asymmetric wells.

Chapters 5 and 6 report on the first demonstration of the validity of Burt's exact envelope function theory for a realistic system. This required the development of a

procedure for extracting envelope functions, as defined by Burt, from bound state wavefunctions generated by the complex band structure EPM approach. These envelope functions were shown to be very similar in form to those calculated by traditional $\mathbf{k} \cdot \mathbf{p}$ methods, except that the derived envelope functions are continuous *and* smooth over all space. However, in the regions around interfaces they vary more rapidly and it is shown that the envelope function gradient follows, closely but smoothly, the usual discontinuity in the gradient of the envelope function generated by an effective mass model. Chapter 6 demonstrates dramatically the inadequacies of the single band edge (SBE) approximation when calculating interband dipole matrix elements and the time dependence of charge oscillations induced by coherent EM excitation. Again, this demonstrates the validity for real systems of predictions made by Burt. Although the SBE approximation reproduces at least 90% of the probability density of a bound state, the contribution to the wavefunction from the other band edge term in the full envelope function expansion contains the symmetry information that is vital for obtaining finite interband dipole matrix elements.

Finally, the chapter addresses the issue of $\mathbf{k} \cdot \mathbf{p}$ differential operator ordering, in collaboration with Paul Stavrinou. The author has generated $\mathbf{k} \cdot \mathbf{p}$ parameters from bulk EPM calculations. These were used by a three band $\mathbf{k} \cdot \mathbf{p}$ calculation of the in-plane energy band structure for valence band quantum wells. Both traditionally accepted 'symmetrised' and new 'ordered' $\mathbf{k} \cdot \mathbf{p}$ Hamiltonians were used and comparison made with equivalent complex band structure EPM calculations. Whilst the $In_{0.53}Ga_{0.47}As/InP$ calculations were inconclusive, those for an ersatz InSb/CdTe well provided striking evidence of the incorrectness of the symmetrised Hamiltonian and the accuracy of the ordered Hamiltonian, both in quantitative and qualitative terms.

To summarise, the complex band structure empirical pseudopotential method is a powerful tool for studying the electronic states of heterostructures. Envelope func-

tion and effective mass models are also often very useful, though great care must be taken both in their range of applicability and in the details of their implementation.

Suggestions for further work

An obvious extension to this work would be to develop the heterostructure software one stage further and calculate calculate actual absorption spectra. Most of the necessary work has already been done, in that the the ability to calculate of the joint density of states and momentum matrix elements between bound states is already in place. This would provide a more useful means of evaluating structures for potential device applications, as well as allowing for more direct comparison with experimental data.

The heterostructure calculations in this thesis have not employed self-consistent techniques. A possibly significant improvement to the techniques employed would be the incorporation of self-consistency and allowance for band bending effects. Such effects would be of greatest importance for the InAs/GaSb/AlSb stepped well, where charge transfer from the GaSb valence band to InAs conduction band takes place with sufficiently wide layers and/or applied electrostatic fields. Such charge transfer may significantly modify the band edge profiles and hence the electronic properties of the system.

Lastly, this thesis is limited to the analysis of bound states only. The author has developed software that treats the heterostructure problem in terms of the scattering at the heterojunction of an incident state, using a complex band structure EPM method very similar to that described in Chapter 3. Whilst this was successful for incident and scattered states within the Γ-valley of each material, X-valley states proved numerical troublesome, with current conservation often being lost. This is due to the lack of convergence of solutions to the pseudo-Hamiltonian for states near the Brillouin zone boundary, when using only 89 reciprocal lattice vectors.

Other workers [120] have worked around this problem by essentially using a non-symmetric set of reciprocal lattice vectors, so that the Fourier representation of the wavefunction is much more accurate in the incident (growth) direction. This has the unfortunate side-effects in terms of loss of symmetry, but it may be that the incorporation into the software of such techniques, possibly in a modified form, would prove useful for the further study of the interesting heterostructures considered in Chapter 4.

Appendix A

Derivation of the pseudo Hamiltonian

To demonstate the derivation (see, for example, [25,28]) of the pseudo Schrödinger equation, (2.5), we first rewrite the energy eigenfunction for valence electrons, (2.4), using Dirac notation,

$$|\psi\rangle = |\varphi\rangle - \sum_c |\psi_c\rangle\langle\psi_c|\varphi\rangle, \tag{A.1}$$

where φ is some function to be determined and ψ_c are core electron energy eigenfunctions. Thus, the valence states, with wavefunction ψ, are automatically orthogonal to the core states. Substituting (A.1) into the Schrödinger equation, (2.1), gives

$$\widehat{H}|\varphi\rangle - \sum_c \widehat{H}|\psi_c\rangle\langle\psi_c|\varphi\rangle = E|\varphi\rangle - \sum_c E|\psi_c\rangle\langle\psi_c|\varphi\rangle. \tag{A.2}$$

Now, the core states are eigenstates of the Hamiltonian, with energy E_c, i.e., $\widehat{H}\,|\psi_c\rangle = E_c\,|\psi_c\rangle$, so, expanding the Hamiltonian in terms of kinetic and potential energy leads

to

$$\widehat{T}|\varphi\rangle + V|\varphi\rangle - \sum_c E_c|\psi_c\rangle\langle\psi_c|\varphi\rangle = E|\varphi\rangle - \sum_c E|\psi_c\rangle\langle\psi_c|\varphi\rangle, \tag{A.3}$$

where $\widehat{T}$ is the kinetic energy operator. Rearranging results in

$$\widehat{T}|\varphi\rangle + \Big(V + \sum_c \big(E - E_c\big)|\psi_c\rangle\langle\psi_c|\Big)|\varphi\rangle = E|\varphi\rangle, \tag{A.4}$$

which is precisely the pseudo-Hamiltonian of (2.5), as required.

Appendix B

The dielectric function and reflectivity

The dielectric function, $\epsilon(\omega)$, which is a function of photon energy, $E = \hbar\omega$, is in general a complex quantity, i.e.,

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega). \tag{B.1}$$

The imaginary part is due to absorption of the incident electromagnetic waves and hence can be related to the excitation of electrons by photon absorption. Hence, combining Fermi's Golden Rule with classical electromagnetic theory leads to a semi-classical expression for the imaginary part of the dielectric function, per unit volume (see, for example, p31 of [25] or Chapter 6 of [3]),

$$\epsilon_2(\omega) = \frac{\pi e^2}{m^2\omega^2\epsilon_0} \sum_{vc} \int |P_{vc}(\mathbf{k})|^2 \frac{d^3\mathbf{k}}{(2\pi)^3} \delta(\Delta E_{\mathbf{k}}), \tag{B.2}$$

where $\omega = E\hbar^{-1}$, $\Delta E_{\mathbf{k}} = (E_{\mathbf{k}} - E)$ and $P_{ij}(\mathbf{k})$ is the momentum matrix element between states in bands i and j at wave vector $\mathbf{k}$. The summation is such that only electrons excited from the valence to conduction bands ($v \rightarrow c$) are included, and

the integration is over a single Brillouin zone.

The Kramers-Kronig relation,

$$\epsilon_1(\omega) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega' \epsilon_2(\omega')}{\omega'^2 - \omega^2} d\omega', \tag{B.3}$$

where P indicates the Cauchy principle value, allows the calculation of the real part of the dielectric function from the imaginary part, The complex refractive index is related to the complex dielectric function by $N^2(\omega) = \epsilon_1 + i\epsilon_2$. Thus, from an initial calculation of ϵ_2 one may obtain the reflectivity,

$$R(\omega) = \left| \frac{N(\omega) - 1}{N(\omega) + 1} \right|^2, \tag{B.4}$$

allowing comparison with a directly measurable quantity.

Appendix C

Single parabolic band approximation

The single parabolic band (SPB) approximation is the most extreme simplification of the effective mass approximation. Each relevant energy band of the electronic band structure of a material is approximated by a parabolic band. Further, in the description of inhomogeneous systems, such as heterostructures or impurities in a bulk crystal, each band is considered in isolation, hence the use of the term 'single'. Since, to a good approximation, zinc-blende materials have a parabolic energy conduction band near the zone centre (Γ-point), and there are no other bands close in energy, the SPB approximation for the calculation of donor states and the properties of conduction band quantum wells and superlattices is expected to be a good one.

The SPB approximation for a homogeneous semiconductor is essentially equivalent to the usual free electron model described in standard quantum mechanics textbooks, but with the free electron mass replaced by an effective mass, m^*. Alternatively, it can be thought of as a $\mathbf{k} \cdot \mathbf{p}$ model that explicitly includes only a single band. Since we are interested in heterostructures that confine the electron in one

direction only (quantum wells, etc), we only consider a one dimensional (1D) model, along the z-direction.

The time independent Schrödinger-like equation for the envelope function multiplying the Bloch periodic part of the relevant band, in a region R, takes the form

$$\left(-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + V_R\right) f_k(z) = E_k f_k(z), \tag{C.1}$$

where V_R is a constant potential within R and E_k is the energy. Solutions of (C.1) have the general form

$$f_k(z) = A_+ e^{ikz} + A_- e^{-ikz}, \tag{C.2}$$

where the A are constants of integration and the wave vector k is given by

$$k = \frac{\sqrt{2m^*(E_k - V_R)}}{\hbar}. \tag{C.3}$$

C.1 Heterostructures

In the application of the SPB approximation to heterostructures, the piece-wise constant potential term within each material region is chosen to recreate the band offsets that form the well. For a single well of depth V_0 and width L, with the real space origin at its centre, the envelope function Schrödinger-like equation is as in (C.1), but with the potential term being

$$V_R = \begin{cases} 0 & |z| \leq L/2, \\ V_0 & |z| > L/2. \end{cases} \tag{C.4}$$

The solutions in each of the three regions are still of the type (C.2) but we must impose the relevant quantum mechanical boundary conditions to obtain the valid energies and determine the coefficients.

If we consider the case of bound states, then we must set one of the coefficients in each barrier region to zero to prevent $f_k(z)$ diverging with increasing or decreasing z. This leaves four unkown constants which are determined by the boundary conditions applied at the well walls, $z = \pm L/2$. This leads to the question of particular interest here: 'What are the appropriate boundary conditions?'

In Section 3.2 it was pointed out that the interface boundary conditions are derived from the form of the Schrödinger equation. However, the derivation of (C.1) is only valid for a piece of homogeneous material and therefore cannot be used to deduce the boundary conditions at a heterojunction. The conservation of current density, which, for a quantum mechanical state ψ, is given by (see, for example, p215 of [1])

$$\mathbf{j} = -\frac{i\hbar}{2m_e}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right). \tag{C.5}$$

Within the EFA, (C.5) becomes, along z,

$$j_z = -\frac{i\hbar}{2m^*}\left(f^*\frac{\partial}{\partial z}f - f\frac{\partial}{\partial z}f^*\right), \tag{C.6}$$

which is often used to suggest a form for the matching conditions for the envelope functions. From (C.6), the current density is continuous at a heterojunction if $f_k(z)$ and $(m^*)^{-1}\partial f_k/\partial z$ are each continuous. However, one can equally match $(m^*)^{a-1}f_k(z)$ and $(m^*)^{-a}\partial f_k/\partial z$, where a make take any value [118]. In other words, the matching conditions appear to be an ad hoc mathematical convenience, without rigorous physical justification. The justification for the continuity of $f_k(z)$ and

$(m^*)^{-1}\partial f_k/\partial z$ requires a microscopic model of the type developed by Burt [17].

Appendix D

EPM parameters and transition energies

D.1 Pseudopotential form factors

Material	v_3^s	v_8^s	v_{11}^s	v_3^a	v_4^a	v_{11}^a
AlSb	−0.209788	0.043316	0.060000	0.060245	0.051541	0.033573
GaSb	−0.184287	0.000939	0.049996	0.075108	0.038286	0.008504
InP	−0.230000	0.014711	0.051069	0.072154	0.053969	0.009583
InAs	−0.231960	−0.005982	0.051448	0.060784	0.048085	0.009996
InSb	−0.218539	0.000000	0.040000	0.058957	0.046679	0.010060
$In_{0.53}Ga_{0.47}As$	−0.235572	0.010084	0.049959	0.053293	0.047697	0.011019
CdTe	−0.260031	−0.000470	0.045000	0.091921	0.072342	0.010295

Table D.1 The pseudopotential form factors, in Rydbergs, used in the EPM calculations (see Section 2.2.2). The values were produced by the fitting procedure described in Section 2.4.1.

For $In_{0.53}Ga_{0.47}As$, rather than obtaining the form factors by interpolating between those of InAs and GaAs (the virtual-crystal approximation), they were found directly by the usual fitting procedure, since enough experimental data exists on

this particular alloy for the present purposes. This approach was chosen because the virtual-crystal method cannot include disorder induced effects, which may be important in determining the properties of the alloy [121].

D.2 Pseudopotential form factor gradients

Material	g_3^s	g_8^s	g_{11}^s	g_3^a	g_4^a	g_{11}^a
AlSb	0.399970	0.123260	−0.039000	−0.087165	−0.087156	−0.000772
GaSb	0.227963	0.151559	0.013918	−0.100002	−0.000013	0.000000
InAs	0.268558	0.172243	−0.033086	−0.088970	−0.046673	−0.000030

Table D.2 The pseudopotential form factor gradients, in Rydbergs/$(2\pi/a)$, used in the EPM calculations, where a is the unstrained lattice parameter of the material (see Section 2.3.3). The values were produced by the fitting procedure described in Section 2.4.3.

D.3 Pseudopotential spin parameters

Material	S_μ
AlSb	0.004146
GaSb	0.003816
InP	0.000540
InAs	0.001827
InSb	0.004676
$In_{0.53}Ga_{0.47}As$	0.001838
CdTe	0.003210

Table D.3 The pseudopotential spin parameters, used in the EPM calculations (see Section 2.2.3). For each material, S_α was set to 1. The values were produced by the fitting procedure described in Section 2.4.2.

D.4 Miscellaneous experimental parameters

Material	a (Å)	c_{11}	c_{12}	c_{44}	Reference
AlSb	6.136	8.769	4.341	4.076	[44]
GaSb	6.095	11.81	5.32	5.94	[44]
InP	5.8688	10.11	5.61	4.56	[122]
InAs	6.0583	8.329	4.526	3.959	[44]
InSb	6.48	6.58	3.49	3.03	[25]
$In_{0.53}Ga_{0.47}As$	5.8688	9.99	4.93	4.89	[122]
CdTe	6.48	—	—	—	[25]

Table D.4 The experimentally determined values of the lattice constant, a, and elastic moduli, c_{ij}, used in the EPM calculations. The data is for room temperature ($\approx$ 300K) and the elastic moduli are given in 10^{11} dyn·cm^{-2}. The lattice parameter of InSb has been chosen to be *exactly* equal to that of CdTe, though its actual value is 6.478Å. Data not used in calculations reported in this work (e.g., InSb c_{ij}) is included for completeness.

D.5 Bulk transition energies

D.5.1 Aluminium antimonide

Transition	Energy (eV) Experimental	Calculated
$E_{g,i}$	1.614	1.613
$E_{g,d}$	2.29	2.295
Δ_0	0.675	0.675
E_L	2.211	2.213
E_1	2.890	2.978
Δ_1	0.42	0.412
a	−5.9	−5.898
b	−1.35	−1.350

Table D.5 Experimental transition energies and deformation potentials for AlSb used in the parameter fitting procedures, along with the final calculated values. The experimental transition energies are taken from [123], except $E_{g,i}$ and $E_{g,d}$, which are the average of values from [123] and [124]. The original values for $E_{g,i}$ are 1.613 and 1.615, and for $E_{g,d}$ are 2.300 and 2.28, respectively. In both cases the data is for 300K/room temperature. The deformation potentials, a and b, are taken from [44] and are for 77K.

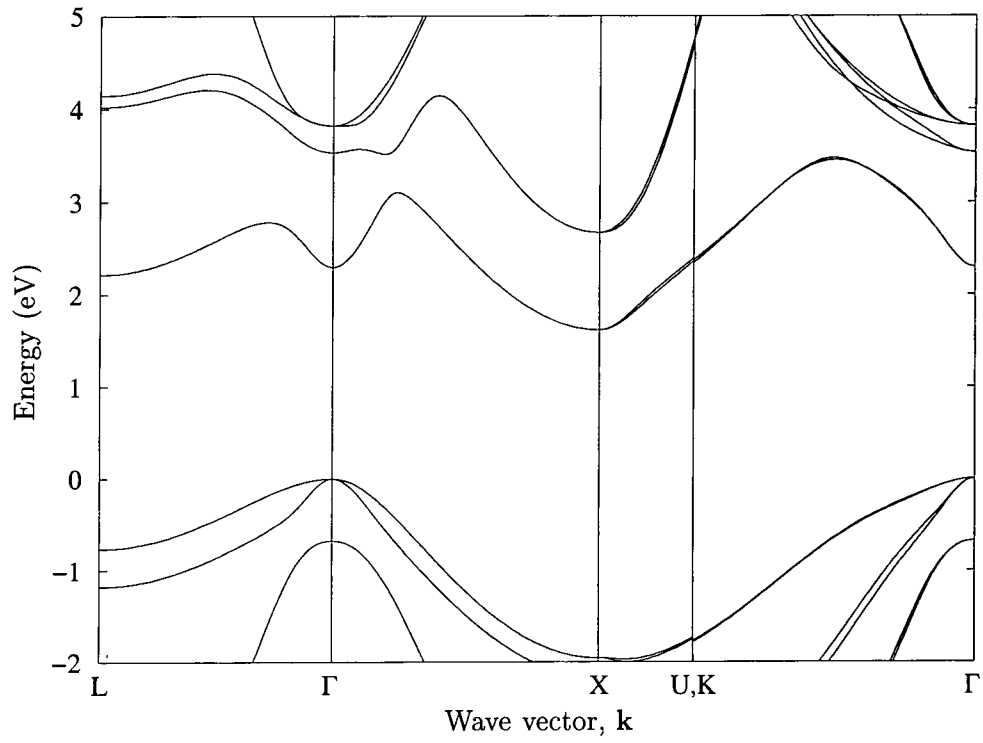


Figure D.1 Bulk band structure of AlSb, calculated using the EPM.

D.5.2 Gallium antimonide

Transition	Energy (eV) Experimental	Calculated
E_g	0.725	0.726
Δ_0	0.76	0.760
E_1	2.055	1.695
Δ_1	0.435	0.455
E_L	0.761	0.890
E_X	1.05	1.046

Table D.6 Experimental transition energies for GaSb used in the parameter fitting procedures, together with the final calculated values. The experimental energies are taken from [123], and are for 300K.

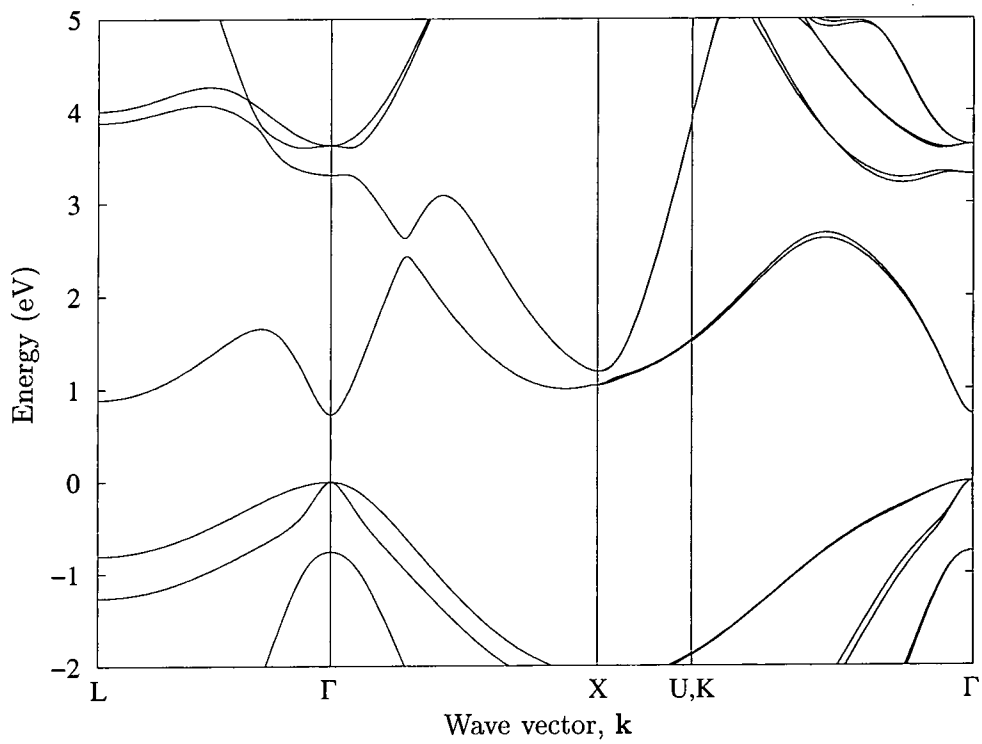


Figure D.2 Bulk band structure of GaSb, calculated using the EPM.

D.5.3 Indium arsenide

Transition	Energy (eV) Experimental	Calculated
E_g	0.356	0.357
Δ_0	0.39	0.391
E_0'	4.6	4.217
E_1	2.50	2.131
E_1'	6.4	5.629
E_2	4.72	3.904
a	−6.0	−6.000
b	−1.8	−1.800

Table D.7 Experimental transition energies and deformation potentials for InAs used in the parameter fitting procedures, together with the final calculated values. The experimental transition energies are taken from [44], and are for 295–300K, except E_0' which is for 80K. The deformation potentials, a and b, are also taken from [44] and are for 300K.

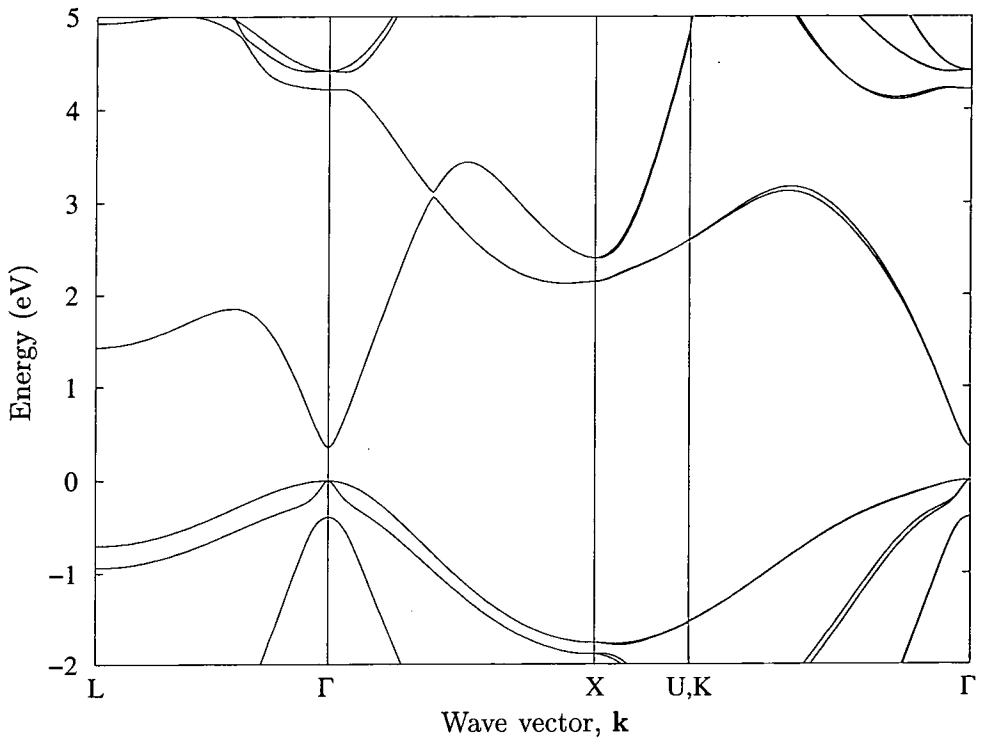


Figure D.3 Bulk band structure of InAs, calculated using the EPM.

D.5.4 Indium phosphide

Transition	Energy (eV) Experimental	Calculated
E_g	1.35	1.355
Δ_0	0.11	0.110
E_0'	4.8	4.382
Δ_0'	0.07	0.048
E_1	3.15	2.640
Δ_1	0.15	0.066
Γ_c–L_c	0.5	0.495
E_2	5.02	4.035
Γ_c–X_c	0.7	0.703

Table D.8 Experimental transition energies for InP used in the parameter fitting procedures, along with the final calculated values. The experimental data is from [44], except for E_g, which is from [122]. All values are for 295–300K, except E_0', which is for 77K.

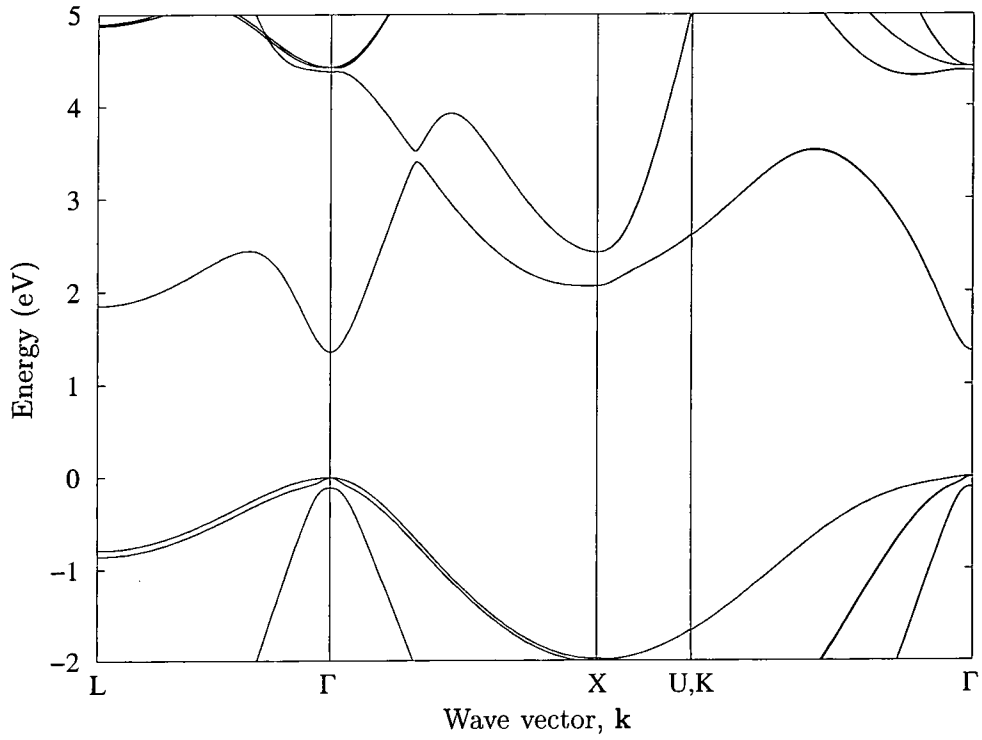


Figure D.4 Bulk band structure of InP, calculated using the EPM.

D.5.5 Indium antimonide

Transition	Energy (eV) Experimental	Calculated
E_g	0.180	0.180
Δ_0	0.98	0.980
E_0'	3.16	3.534
Δ_0'	0.33	0.480
E_1	1.88	1.903
E_1'	5.25	4.890
Δ_1	0.50	0.597
E_2	4.08	3.637

Table D.9 Experimental transition energies for InSb used in the parameter fitting procedures, along with the final calculated values. The experimental transition energies are taken from [44], and are for 300K.

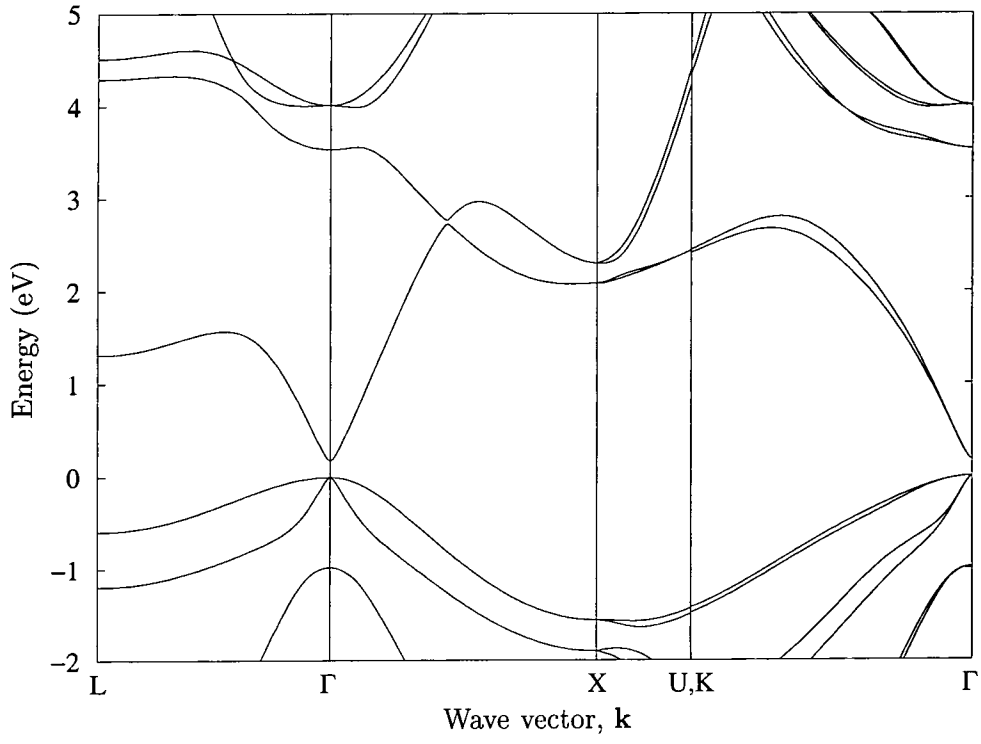


Figure D.5 Bulk band structure of InSb, calculated using the EPM.

D.5.6 $In_{0.53}Ga_{0.47}As$

Transition	Energy (eV) Experimental	Calculated
E_g	0.75	0.753
Δ_0	0.37	0.370
E_1	2.65	2.232
Δ_1	0.24	0.224
E_L	1.39	1.393
E_X	1.68	1.683

Table D.10 Experimental transition energies for $In_{0.53}Ga_{0.47}As$ used in the parameter fitting procedures, together with the final calculated values. The experimental transition energies are taken from [122], although the value for E_X is calculated using $E_X^{(InAs)} = 2.14$. This InAs energy corresponds to the present calculations. All data is for room temperature.

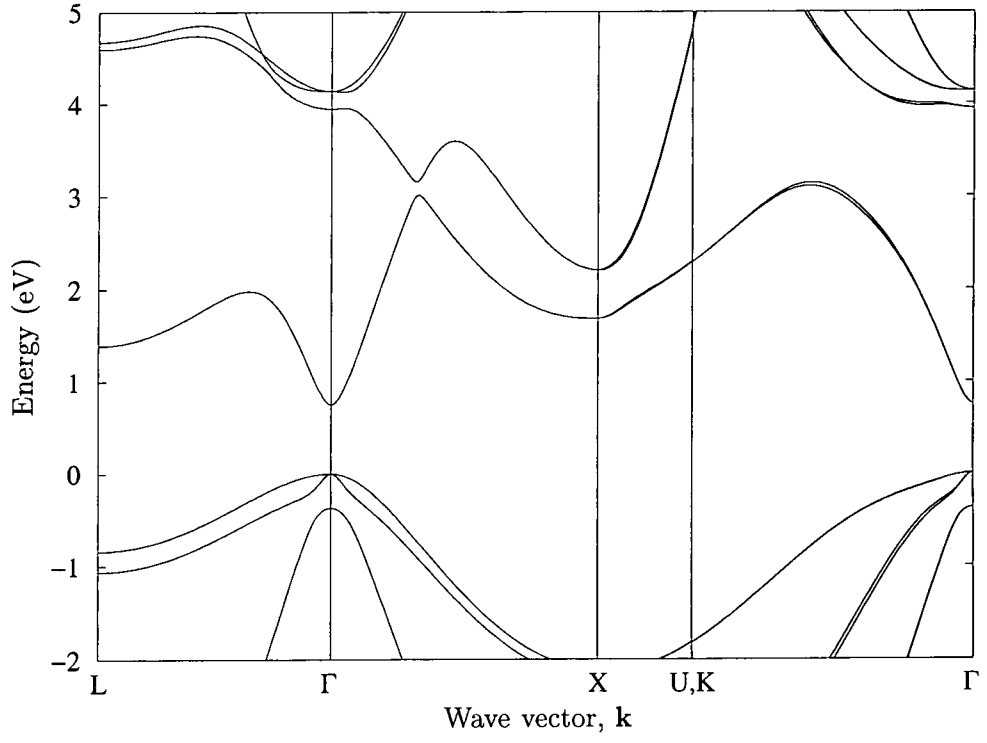


Figure D.6 Bulk band structure of $In_{0.53}Ga_{0.47}As$, calculated using the EPM.

D.5.7 Cadmium telluride

Transition	Energy (eV) Experimental	Calculated
E_g	1.44	1.440
Δ_0	0.84	0.840
E'_0	5.16	5.344
E_1	3.07	3.325
Δ_1	0.56	0.496
E_2	5.40	5.100

Table D.11 Experimental transition energies for CdTe used in the parameter fitting procedures, together with the final calculated values. The experimental transition energies are taken from [125], and are for 300K.

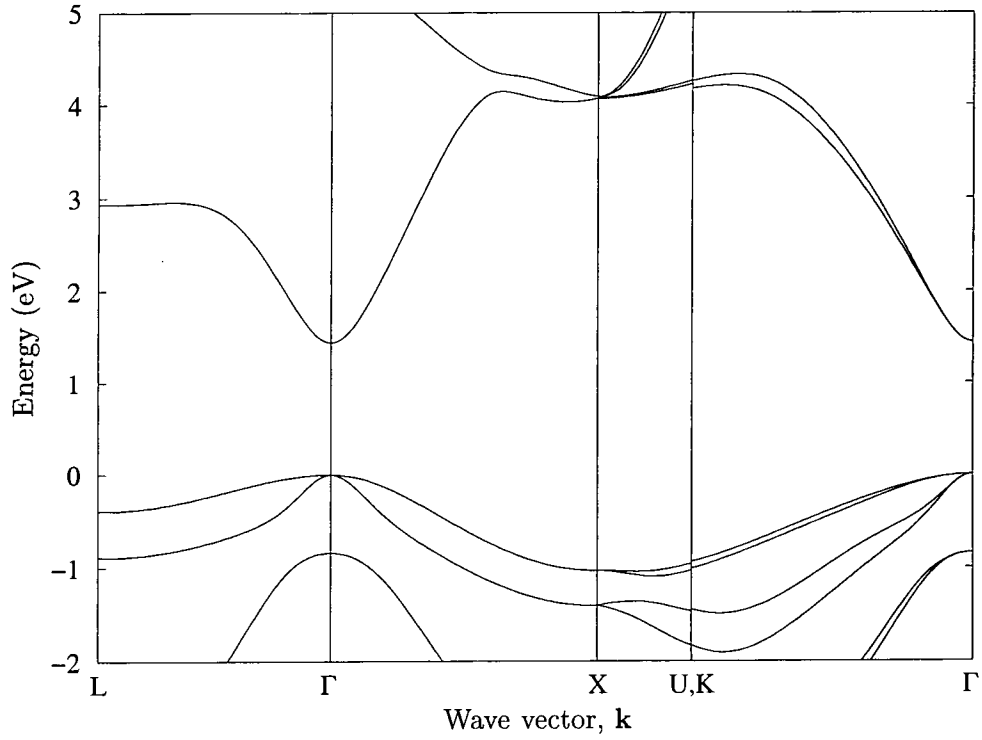


Figure D.7 Bulk band structure of CdTe, calculated using the EPM.

D.6 Complex band structure

D.6.1 Aluminium antimonide

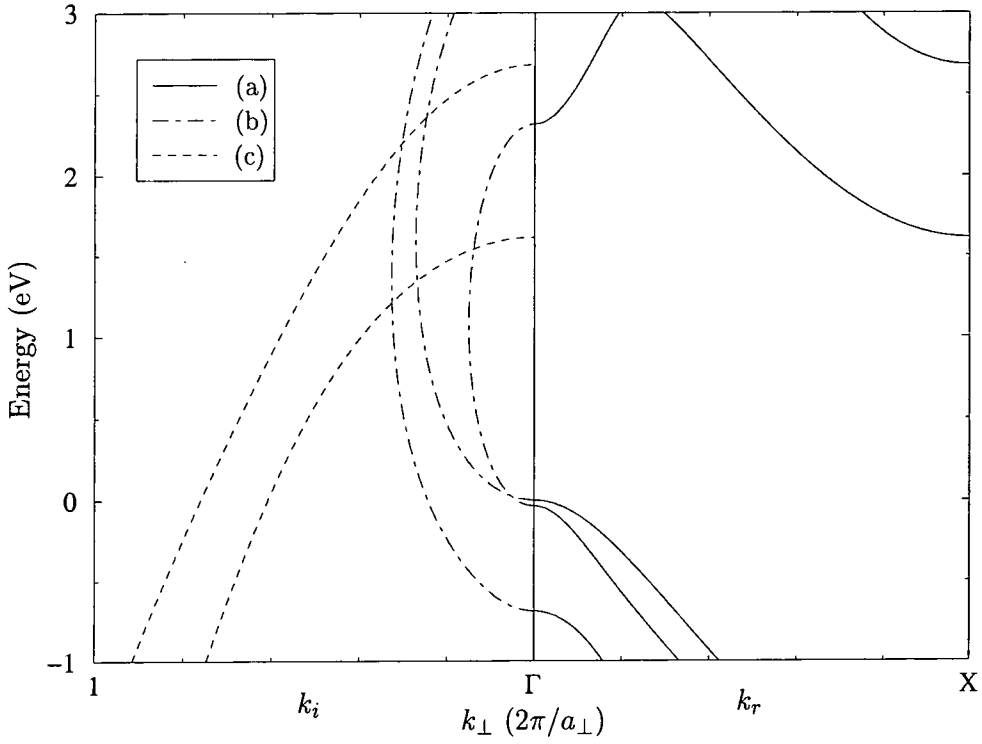


Figure D.8 Strained complex band structure of AlSb. The strain is consistent with dislocation free growth on GaSb, leading to negative strain in the in-plane direction and hence positive strain in the perpendicular (growth) direction. Thus, along the growth direction, the heavy hole band is raised above the light hole band around the zone centre [37]. The energy bands are labelled as follows: (a) purely real k, (b) purely imaginary k, (c) complex k with real part at X.

D.6.2 Gallium antimonide

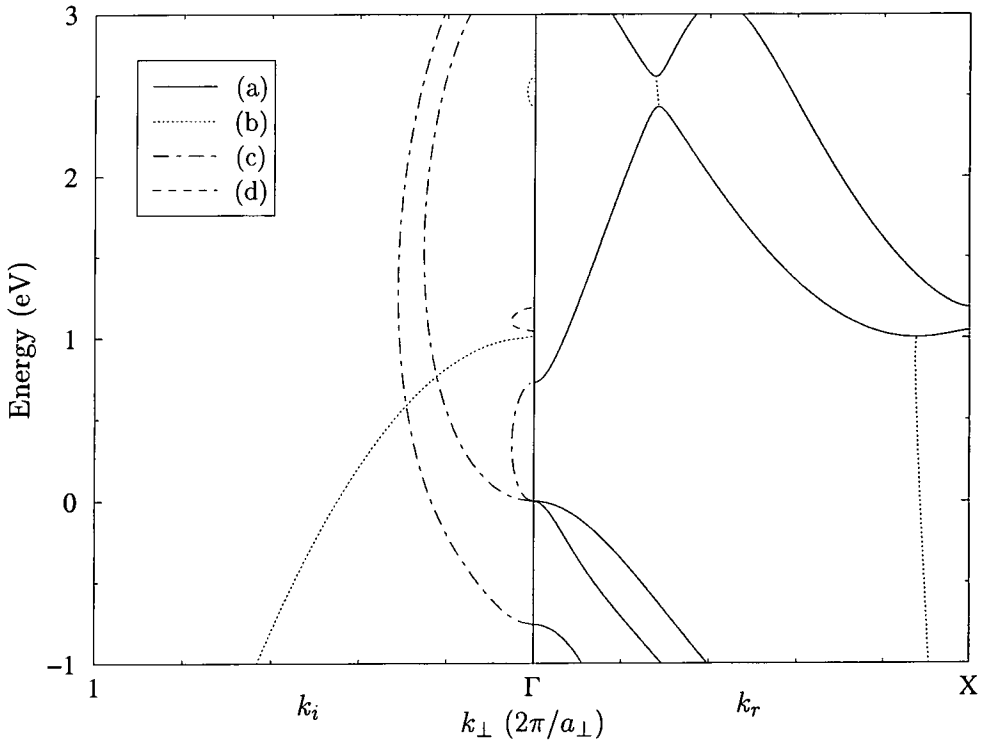


Figure D.9 Complex band structure of GaSb. The energy bands are labelled as follows: (a) purely real k, (b) complex k, (c) purely imaginary k, (d) complex k with real part at X.

D.6.3 Indium arsenide

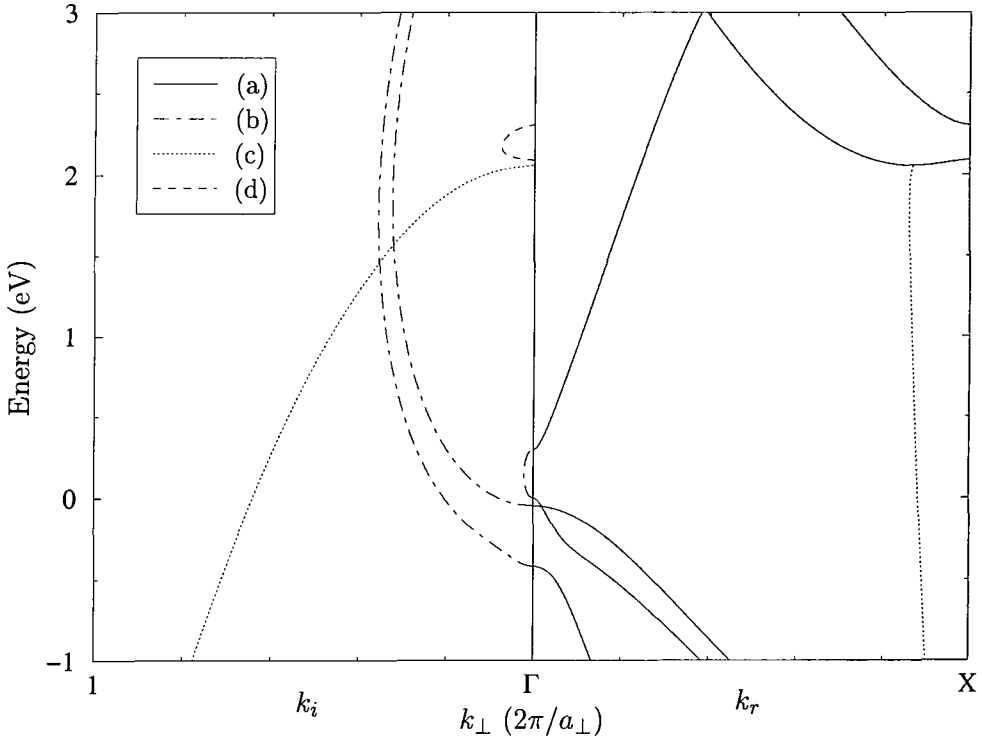


Figure D.10 Strained complex band structure of InAs. The strain is consistent with dislocation free growth on GaSb, leading to positive strain in the in-plane direction and hence negative strain in the perpendicular (growth) direction. Thus, in the growth direction, the light hole band is raised above the heavy hole band around the zone centre [37]. The energy bands are labelled as follows: (a) purely real k, (b) purely imaginary k, (c) complex k, (d) complex k with real part at X.

D.7 Heterojunction band line-ups

Unfortunately, values for heterojunction energy band line-ups are not always as well established as would be preferable. Much experimental and theoretical effort has been put in to the determination of numerous band line-ups, however, theoretical predictions for the same heterojunction vary considerably, so the data used in this work is taken from experiment, although sometimes experimental values are also in poor agreement. Yu *et al.* [62] have published a review of this topic and collated the various data available at the time, commenting on the validity or otherwise of the data where a consensus has emerged. In the following sections, each system modelled in this thesis is discussed in a little detail, specifying any additional references.

D.7.1 InAs/GaSb/AlSb system

Materials	ΔE_v (eV)	Reference
InAs/GaSb	0.50	[126]
	0.56	[127]
	0.56	[128]
	0.51 ± 0.1	[129]
	0.52	[130]
AlSb/GaSb	0.04–0.08	[131]
	> 0.267	[132]
	0.40 ± 0.15	[133]
	0.45 ± 0.08	[134]
	0.35	[135]
	0.39 ± 0.07	[136]
InAs/AlSb	0.16	In-Sb [137]
	0.17	In-Sb [137]
	0.22	Al-As [137]
	0.15	Al-As [137]

Table D.12 Experimentally determined band offsets for the InAs/GaSb/AlSb system, taken mostly from Yu *et al.* [62]. The GaSb valence band maximum lies above that of InAs and AlSb. The AlSb valence band maximum lies above that of InAs. The InAs/AlSb results are for InAs grown on relaxed AlSb. The InAs/AlSb entries also indicate the interface type formed during growth.

There is only one reference for InAs/AlSb experimental data. However, recent the-

Materials	ΔE_v (eV) Used	Exp.
InAs/GaSb	0.51	0.53
AlSb/GaSb	0.40	0.40
InAs/AlSb	0.11	0.18

Table D.13 Offset values used in the calculations and average experimental values.

oretical calculations [138] predict an offset of $\Delta E_v = 0.19$eV (AlSb above InAs), for both interface types, which is consistent with the experimental result.

The average value of ΔE_v for the InAs/GaSb heterojunction is 0.53eV, for the AlSb/GaSb heterojunction it is 0.40eV (excluding the first two results), and for InAs/AlSb it is 0.18eV. While it is desirable to use these values in the EPM calculations, in a heterostructure consisting of all three materials, such as that described in Section 4.3, they are not quite consistent. Therefore, one or more of the values must be modified. Additional experimental data is available to assist in this process. The band overlap between InAs and GaSb, where the InAs conduction band minimum lies below the GaSb valence band maximum, has recently been measured as 0.146eV at 300K [139]. In the present calculations, the InAs band gap has been fitted to experiment (including strain effects) with a value of 0.299eV. For a given InAs band gap, varying ΔE_v also varies the band overlap, and so a balance between the two must be struck.

The values actually used are given in Table D.13. These are in satisfactory agreement, with the exception of the InAs/AlSb offset, and lead to a InAs/GaSb band overlap of 0.211eV. Increasing the InAs offset relative to both GaSb and AlSb seems desirable, but this would increase the band overlap with GaSb, which is already possibly too large, according to [139]. On the other hand, a value for the InAs/AlSb ΔE_v of 0.1eV has recently been quoted [140,141], hence the offsets used in the calculations provide a reasonable compromise to the available experimental

data.

D.7.2 $In_{0.53}Ga_{0.47}As/InP$ system

Experimental determination of the conduction band offset for the $In_{0.53}Ga_{0.47}As/InP$ system typically give a value of approximately $0.4\Delta E_g$ [62], where ΔE_g is the difference in the fundamental band gaps of the two materials. A value of 0.38eV [122] is used for the valence band offset in the present work. This is equivalent to a conduction band offset of $0.37\Delta E_g$.

D.7.3 InSb/CdTe system

Published data for the CdTe/InSb system is extremely sparse, with only one experimental and one theoretical offset available. Fortunately, these are in very good agreement, suggesting that the data is reasonably reliable. Experiment suggests a valence band offset of 0.87 ± 0.10eV [142] for both (100) and (110) interfaces, whilst an ab initio pseudopotential calculation for (110) interfaces predicts a value of 0.88eV [143]. This work uses a value of 0.87eV.

Bibliography

[1] R. L. Liboff, *Introductory Quantum Mechanics, 2nd edition*, Addison-Wesley, New York (1992)

[2] L. Esaki and R. Tsu, IBM J. Res. Devel. **14**, 61 (1970)

[3] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, Springer-Verlag, Berlin (1996)

[4] J. R. Arthur, J. Appl. Phys. **39**, 4032 (1968)

[5] H. M. Manasevit, Appl. Phys. Lett. **12**, 136 (1968)

[6] B. Brar, H. Kroemer, J. Ibbetson and J. H. English, Appl. Phys. Lett. **62**, 3303 (1993)

[7] J. Spitzer, A. Höpner, M. Kuball, M. Cardona, B. Jenichen, H. Neuroth, B. Brar and H. Kroemer, J. Appl. Phys. **77**, 811 (1995)

[8] L. Esaki and L. L. Chang, Phys. Rev. Lett. **33**, 495 (1974)

[9] D. Mukherji and B. R. Nag, Phys. Rev. B **12**, 4338 (1975)

[10] T. B. Boykin, Appl. Phys. Lett. **64**, 1529 (1994)

[11] Y.-C. Chang and J. N. Schulman, Phys. Rev. B **31**, 2056 (1985)

[12] Y.-C. Chang and J. N. Schulman, Phys. Rev. B **31**, 2069 (1985)

[13] D. M. Wood and A. Zunger, Phys. Rev. B **53**, 7949 (1996)

[14] B. A. Foreman, Phys. Rev. B **52**, 12241 (1995)

[15] A. T. Meney, B. Gonul and E. P. O'Reilly, Phys. Rev. B **50**, 10893 (1994)

[16] B. A. Foreman, Phys. Rev. B **48**, 4964 (1993)

[17] M. G. Burt, J. Phys.: Condens. Matt. **4**, 6651 (1992)

[18] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*, John Wiley & Sons, New York (1988)

[19] M. J. Shaw, P. R. Briddon and M. Jaros, Phys. Rev. B **52**, 16341 (1995)

[20] R. I. Taylor, M. G. Burt and R. A. Abram, Superlatt. Microstruct. **3**, 63 (1987)

[21] S. Brand and D. T. Hughes, Semicond. Sci. Technol. **2**, 607 (1987)

[22] D. L. Smith and C. Mailhiot, Rev. Mod. Phys. **62**, 173 (1990)

[23] M. C. Payne *et al.*, Rev. Mod. Phys. **64**, 1045 (1992)

[24] W. A. Harrison, *Electronic Struture and the Properties of Solids*, W. H. Freeman and Company, San Francisco (1980)

[25] M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, Springer-Verlag, Berlin (1988)

[26] C. Herring, Phys. Rev. **57**, 1169 (1940)

[27] M. L. Cohen and V. Heine, Solid State Physics **24**, 37 (1970)

[28] W. A. Harrison, *Pseudopotentials in the Theory of Metals*, W. A. Benjamin, New York (1966)

[29] J. C. Phillips and L. Kleinman, Phys. Rev. **116**, 287 (1959)

[30] M. L. Cohen and V. Heine, Phys. Rev. **122**, 1821 (1961)

[31] J. R. Chelikowsky and M. L. Cohen, Phys. Rev. B **14**, 556 (1976)

[32] M. L. Cohen and T. K. Bergstresser, Phys. Rev. **141**, 789 (1966)

[33] J. J. Sakurai, *Advanced Quantum Mechanics*, Addison-Wesley, New York (1967)

[34] A. O. E. Animalu, Phil. Mag. **13**, 53 (1966)

[35] G. Weisz, Phys. Rev. B **149**, 504 (1966)

[36] S. Bloom and T. K. Bergstresser, Solid State Commun. **6**, 465 (1968)

[37] E. P. O'Reilly, Semicond. Sci. Technol. **4**, 121 (1989)

[38] P. Voisin, Quantum Wells and Superlattices in Optoelectronic Devices and Integrated Optics, SPIE **861**, 88 (1988)

[39] I. J. Fritz, S. T. Picraux, L. R. Dawson, T. J. Drummond, W. D. Laidig and N. G. Anderson, Appl. Phys. Lett. **46**, 967 (1985)

[40] T. J. Andersson, Z. G. Chen, V. D. Kulakovskii, A. Uddin and J. T. Vallin, Appl. Phys. Lett. **51**, 752 (1987)

[41] R. People and J. C. Bean, Appl. Phys. Lett. **47**, 322 (1985)

[42] E. Kasper, Surface Sci. **174**, 630 (1986)

[43] J. F. Nye, *Physical Properties of Crystals*, Oxford University Press, Oxford (1957)

[44] O. Madelung, ed., *Landolt-Börnstein, Numerical Data and Functional Relationships in Science and Technology, Volume 17a*, Springer-Verlag, Berlin (1982)

[45] D. T. Hughes, *Electronic States in Semiconductor Superlattices and Quantum Wells*, Ph.D. thesis, University of Durham (1989)

[46] F. Bassani and D. Brust, Phys. Rev. **131**, 1524 (1963)

[47] Y. F. Tsay, S. S. Mitra and B. Bendow, Phys. Rev. B **10**, 1476 (1974)

[48] D. O. Hall, *Calculations of the electronic structure and optical properties of strained II–VI superlattices*, Ph.D. thesis, University of Durham (1991)

[49] L. R. Saravia and D. Brust, Phys. Rev. **178**, 1240 (1969)

[50] B. K. Ridley, *Quantum Processes in Semiconductors*, Clarendon Press, Oxford (1982)

[51] K. C. Pandey and J. C. Phillips, Phys. Rev. B **9**, 1552 (1974)

[52] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Saunders College Publishing (1976)

[53] Y.-C. Chang and J. N. Schulman, Phys. Rev. B **25**, 3975 (1982)

[54] V. Heine, Proc. Phys. Soc. **81**, 300 (1963)

[55] J. P. Cuypers, *Scattering of Electrons at Heterostructure Interfaces*, Ph.D. thesis, Technische Universiteit Eindhoven (1992)

[56] L. van Hove, Phys. Rev. **89**, 1189 (1953)

[57] W. Kohn, Phys. Rev. **115**, 809 (1959)

[58] E. I. Blount, Solid State Physics **13**, 362 (1962)

[59] V. Heine, Surface Sci. **2**, 1 (1964)

[60] R. O. Jones, Proc. Phys. Soc. **89**, 443 (1966)

[61] Y.-C. Chang, Phys. Rev. B **25**, 605 (1982)

[62] E. T. Yu, J. O. McCaldin and T. C. McGill, Solid State Physics **46**, 1 (1992)

[63] H. Kroemer, Surface Sci. **132**, 543 (1983)

[64] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics (Non-relativistic Theory)*, Pergamon Press, Oxford (1977)

[65] J. J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley, New York (1985)

[66] B. H. Bransden and C. J. Joachain, *Introduction to Quantum Mechanics*, Longman, London (1989)

[67] A. C. G. Wood, *Strain effects in semiconductor quantum wells*, Ph.D. thesis, University of Durham (1990)

[68] B. I. Bleaney and B. Bleaney, *Electricity and Magnetism Volume 1, 3rd edition*, Oxford University Press, Oxford (1989)

[69] D. H. Chow, H. L. Dunlap, W. Williamson, III, S. Enquist, B. K. Gilbert, S. Subramaniam, P.-M. Lei and G. H. Bernstein, IEEE Electron Device Lett. **17**, 69 (1996)

[70] R. J. Warburton, C. Gauer, A. Wixforth, J. P. Kotthaus, B. Brar and H. Kroemer, Phys. Rev. B **53**, 7903 (1996)

[71] F. Fuchs, J. Schmitz, J. D. Ralston and P. Koidl, Phys. Rev. B **49**, 13638 (1994)

[72] T. C. McGill and D. A. Collins, Semicond. Sci. Technol. **8**, S1 (1993)

[73] D. M. Symons, M. Lakrimi, R. J. Warburton, R. J. Nicholas, N. J. Mason, P. J. Walker, M. I. Eremets and G. Hill, Phys. Rev. B **49**, 16614 (1994)

[74] J. Genoe, K. Fobelets, C. van Hoof and G. Borghs, Phys. Rev. B **52**, 14025 (1995)

[75] M. H. Liu, Y. H. Wang and M. P. Houng, Jpn. J. Appl. Phys. **33**, 4506 (1994)

[76] J. R. Söderström, D. H. Chow and T. C. McGill, Appl. Phys. Lett. **55**, 1094 (1989)

[77] Q. Du, J. Alperin and W. I. Wang, Appl. Phys. Lett. **67**, 2218 (1995)

[78] J. R. Meyer, C. A. Hoffman, F. J. Bartoli and L. R. Ram-Mohan, Appl. Phys. Lett. **67**, 2756 (1995)

[79] H. Xie, W. I. Wang and J. R. Meyer, J. Appl. Phys. **76**, 92 (1994)

[80] R. Q. Yang, Superlatt. Microstruct. **17**, 77 (1995)

[81] S. A. Hosseini, M. J. Shaw and M. Jaros, Phys. Rev. B **53**, 6988 (1996)

[82] H. Ohno, L. Esaki and E. E. Mendez, Appl. Phys. Lett. **60**, 3153 (1992)

[83] A. Yariv, *Optical Electronics*, Oxford University Press, Oxford (1991)

[84] R. J. Warburton, K. Weilhammer, J. P. Kotthaus, M. Thomas and H. Kroemer, Phys. Rev. Lett. **80**, 2185 (1998)

[85] H. C. Liu, M. Buchanan and Z. R. Wasilewski, Appl. Phys. Lett. **72**, 1682 (1998)

[86] W. Batty and K. A. Shore, IEE Proc.-Optoelectron. **145**, 21 (1998)

[87] G. Bastard, E. E. Mendez, L. L. Chang and L. Esaki, Phys. Rev. B **28**, 3241 (1983)

[88] A. R. Rundell, G. P. Srivastava, J. C. Inkson, E. A. Johnson and A. MacKinnon, Physica B **251**, 710 (1998)

[89] M. J. Yang, C. H. Yang, B. R. Bennett and B. V. Shanabrook, Phys. Rev. Lett. **78**, 4613 (1997)

[90] J. J. Quinn and J. J. Quinn, Surface Sci. **362**, 930 (1996)

[91] K. B. Wong, G. K. A. Gopir, J. P. Hagon and M. Jaros, Semicond. Sci. Technol. **9**, 2210 (1994)

[92] R. Eppenga, M. F. H. Schuurmans and S. Colak, Phys. Rev. B **36**, 1554 (1987)

[93] M. F. H. Schuurmans and G. W. 't Hooft, Phys. Rev. B **31**, 8041 (1985)

[94] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957)

[95] E. O. Kane, Semiconductors and Semimetals **1**, 75 (1966)

[96] M. G. Burt, Semicond. Sci. Technol. **2**, 460 (1987)

[97] M. G. Burt, Semicond. Sci. Technol. **3**, 739 (1988)

[98] R. van Dalen and P. N. Stavrinou, Semicond. Sci. Technol. **13**, 11 (1998)

[99] P. N. Stavrinou and R. van Dalen, Phys. Rev. B **55**, 15456 (1997)

[100] R. Gómez-Alcalá, F. J. Fraile-Peláez, A. Espiñeira and X. Piñeiro, Superlatt. Microstruct. **20**, 35 (1996)

[101] B. A. Foreman, Phys. Rev. Lett. **81**, 425 (1998)

[102] J. P. Cuypers and W. van Haeringen, Physica B **168**, 58 (1991)

[103] J. P. Cuypers and W. van Haeringen, J. Phys.: Condens. Matt. **4**, 2587 (1992)

[104] R. A. Abram and M. Jaros, eds., *Band Structure Engineering in Semiconductor Microstructures, NATO ASI Series B, Volume 189*, Plenum Press, New York (1988)

[105] H. J. Weaver, *Applications of Discrete and Continuous Fourier Analysis*, John Wiley & Sons, New York (1983)

[106] M. G. Burt, J. Phys.: Condens. Matt. **5**, 4091 (1993)

[107] M. G. Burt, Superlatt. Microstruct. **17**, 335 (1995)

[108] M. G. Burt, Semicond. Sci. Technol. **10**, 412 (1995)

[109] G. Bastard, Phys. Rev. B **25**, 7584 (1982)

[110] M. Altarelli, Phys. Rev. B **28**, 842 (1983)

[111] C. Kittel, *Introduction to Solid State Physics, 6th edition*, John Wiley & Sons, New York (1986)

[112] E. Rosencher, A. Fiore, B. Vinter, V. Berger, P. Bois and J. Nagle, Science **271**, 168 (1996)

[113] C. Weisbuch and B. Vinter, *Quantum Semiconductor Structures: Fundamentals and Applications*, Academic Press, Boston (1991)

[114] H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductor*, World Scientific, Singapore (1990)

[115] L. C. West and S. J. Eglash, Appl. Phys. Lett. **46**, 1156 (1985)

[116] P. Löwdin, J. Chem. Phys. **19**, 1396 (1951)

[117] G. Bastard, Phys. Rev. B **24**, 5693 (1981)

[118] R. A. Morrow and K. R. Brownstein, Phys. Rev. B **30**, 678 (1984)

[119] J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955)

[120] A. C. Marsh and J. C. Inkson, Semicond. Sci. Technol. **1**, 285 (1986)

[121] O. Berolo, J. C. Woolley and J. A. van Vechten, Phys. Rev. B **8**, 3794 (1973)

[122] S. Adachi, *Physical Properties of III-V Semiconductor Compounds*, John Wiley & Sons, New York (1992)

[123] C. Alibert, A. Joullié and A. M. Joullié, Phys. Rev. B **27**, 4946 (1986)

[124] K. Strössner, S. Ves, C. K. Kim and M. Cardona, Phys. Rev. B **33**, 4044 (1986)

[125] M. Cardona and D. L. Greenaway, Phys. Rev. **131**, 98 (1963)

[126] H. Sakaki, L. L. Chang, R. Ludeke, C.-A. Chang, G. A. Sai-Halasz and L. Esaki, Appl. Phys. Lett. **31**, 211 (1977)

[127] G. A. Sai-Halasz, L. L. Chang, J.-M. Welter, C.-A. Chang and L. Esaki, Solid State Commun. **27**, 935 (1978)

[128] L. M. Claessen, J. C. Maan, M. Altarelli, P. Wyder, L. L. Chang and L. Esaki, Phys. Rev. Lett. **57**, 2556 (1986)

[129] G. J. Gualtieri, G. P. Schwartz, R. G. Nuzzo, R. J. Malik and J. F. Walker, J. Appl. Phys. **61**, 5337 (1987)

[130] J. R. Waterman, B. V. Shanabrook, R. J. Wagner, M. J. Yang, J. L. Davis and J. P. Omaggio, Semicond. Sci. Technol. **8**, S106 (1993)

[131] P. Voisin, C. Delalande, M. Voos, L. L. Chang, A. Segmuller, C.-A. Chang and L. Esaki, Phys. Rev. B **30**, 2276 (1984)

[132] C. Tejedor, J. M. Calleja, F. Meseguer, E. E. Mendez, C.-A. Chang and L. Esaki, Phys. Rev. B **32**, 5303 (1985)

[133] G. J. Gualtieri, G. P. Schwartz, R. G. Nuzzo and W. A. Sunder, Appl. Phys. Lett. **49**, 1037 (1986)

[134] J. Menéndez, A. Pinczuk, D. J. Werder, J. P. Valladares, T. H. Chiu and W. T. Tsang, Solid State Commun. **61**, 703 (1987)

[135] U. Cebulla, G. Tränkle, U. Ziem, A. Forchel, G. Griffiths, H. Kroemer and S. Subbanna, Phys. Rev. B **37**, 6278 (1988)

[136] E. T. Yu, M. C. Phillips, D. H. Chow, D. A. Collins, M. W. Wang, J. O. McCaldin and T. C. McGill, Phys. Rev. B **46**, 13379 (1992)

[137] J. R. Waldrop, G. J. Sullivan, R. W. Grant, E. A. Kraut and W. A. Harrison, J. Vac. Sci. Technol. B **10**, 1773 (1992)

[138] R. G. Dandrea and C. B. Duke, Appl. Phys. Lett. **63**, 1795 (1993)

[139] D. M. Symons, M. Lakrimi, M. van der Burgt, T. A. Vaughan, R. J. Nicholas, N. J. Mason and P. J. Walker, Phys. Rev. B **51**, 1729 (1995)

[140] Y. W. Chen, H. S. Li and K. L. Wang, Superlatt. Microstruct. **14**, 137 (1993)

[141] C. Gauer, J. Scriba, A. Wixford, J. P. Kotthaus, C. Nguyen, G. Tuttle, J. H. English and H. Kroemer, Semicond. Sci. Technol. **8**, S137 (1993)

[142] K. J. Mackey, P. M. C. Allen, W. G. Herrenden-Harker, R. H. Williams, C. R. Whithouse and G. M. Williams, Appl. Phys. Lett. **49**, 354 (1986)

[143] A. Qteish and R. J. Needs, Phys. Rev. B **47**, 3714 (1993)

Author index

Numbers in italics indicate pages in the bibliography.